CANADIAN JOURNAL OF PSYCHOLOGY

VOLUME 13, No. 4 DECEMBER, 1959

PUBLISHED FOR THE
CANADIAN PSYCHOLOGICAL ASSOCIATION BY THE
UNIVERSITY OF TORONTO PRESS

AUTHORIZED AS SECOND-CLASS MAIL, POST-OFFICE DEPARTMENT, OTTAWA

CANADIAN JOURNAL OF PSYCHOLOGY

Editor: JULIAN BLACKBURN *Assistant Editor:* P. H. R. JAMES

THE CANADIAN JOURNAL OF PSYCHOLOGY is published quarterly in March, June, September, and December. Annual subscription, $6.00, single number, $1.50.

Subscriptions. Orders and correspondence regarding subscriptions, change of address, and purchase of back numbers should be sent to:

THE SECRETARY-TREASURER, CANADIAN PSYCHOLOGICAL ASSOCIATION

Box 31, Postal Station D., Ottawa, Ontario

Contributions. Original manuscripts and correspondence on editorial matters should be sent to:

THE EDITOR, CANADIAN JOURNAL OF PSYCHOLOGY

Queen's University, Kingston, Ontario

Information for contributors. The *Journal* publishes experimental and theoretical articles in all recognized fields of psychology. Contributors need not be members of the CPA or residents of Canada. Manuscripts should be submitted in duplicate, they should be double-spaced throughout, and they should follow standard practice as regards tables, figures, etc. Authors are asked to adopt the current practice in American and British journals of citing references by date, for example, Smith (1956), instead of by number. "Immediate publication" (i.e. in the next issue to go to press) can be arranged for authors willing to pay the extra costs involved.

CANADIAN PSYCHOLOGICAL ASSOCIATION, 1959–60

President: G. H. TURNER, London; *Past President:* DALBIR BINDRA, Montreal; *President Elect:* R. B. BROMILEY, Toronto; *Secretary-Treasurer:* F. R. WAKE, Ottawa; *Employment Bureau:* MISS MARJORIE FAHRIG, 68 Tower Road, Ottawa 5.

1959 Annual Meeting

Kingston, Ontario—June 2, 3, 4

The Canadian Psychological Association also publishes *The Canadian Psychologist*, which is distributed to members only. *Editor:* W. R. N. BLAIR, 462 Melbourne Ave., Ottawa, Ont.

HEART RATE CHANGES DURING AVOIDANCE LEARNING IN DOGS

A. H. BLACK[1]

McMaster University

MUCH OF THE PSYCHOLOGICAL RESEARCH on heart rate has been concerned with "emotional" behaviour. Such research can be classified into two categories. First is the differential analysis of patterns of skeletal and autonomic behaviour that are elicited by varying types of so-called "emotional" unconditioned stimuli. Ax (1), for example, has shown that stimulus situations which are "fear-provoking" produce a pattern of response different from those which are "anger-provoking." In the second type of experiment, the development of learned or conditioned emotional responding has been the main concern, and here, classical conditioning methods have provided most of the data. For example, Gantt (5) has carried out an extensive programme on the analysis of conditioned cardiac responses using both aversive and positive unconditioned stimuli.

In the present experiment, heart rate responding was measured during the acquisition and extinction of an instrumental avoidance response in dogs. While the avoidance training procedure differs considerably from that used in classical conditioning, gross behavioural observations on responses such as defecation, panting, barking, and trembling have shown that regular changes in emotional behaviour occur during avoidance learning in dogs (13). Furthermore, Bersh, Notterman, and Schoenfeld (2) have shown some correlation between avoidance and heart rate behaviour in human subjects when the avoidance procedure was instituted following classical conditioning. It was hoped that the measurement of heart rate during avoidance learning in this experiment would provide a quantitative index of emotional responding and supply empirical data for a comparison of the effects of classical conditioning and avoidance training procedures on heart rate.

The results of this experiment would seem to be relevant to theories of avoidance learning such as that proposed by Mowrer (9) who assumed

[1]This study was supported by funds from the Rockefeller Foundation, and was facilitated by the Laboratory of Social Relations, Harvard University. Part of the study was done while the author was an SSRC Fellow.

The author is indebted to Dr. Richard L. Solomon for his continued assistance and advice, and to Mr. Malcolm Westcott for aid in running the experiment.

that conditioned autonomic responses play an important role in avoidance learning as mediators of motivation and reinforcement. However, there is some difficulty in interpreting heart rate data theoretically because of the apparent control of heart rate by variables such as skeletal activity and respiration, in addition to the direct influence of the conditioning procedures themselves. Smith (11) has argued that all so-called conditioned heart rate responses are artifacts of concomitant activity. McLeary (7) has provided some data in support of this argument. Peters and Gantt, on the other hand, minimize this difficulty. "Therefore, whether the rate of the heart beat is increased reflexly by muscular tension or whatever the relation between *HR* and muscular tension might be, we think that muscular tension has little to do with the great increases in *HR* that go with conditional reflexes since these reflexes are accompanied by a very small degree of muscular tension" (10, p. 40). Since changes in skeletal behaviour regularly occur during the course of avoidance learning, the interpretation of positive results to support a theory such as Mowrer's is necessarily ambiguous. Regular changes in heart rate which parallel regular changes in avoidance learning may possibly be reflex changes in heart rate produced by the occurrence of the avoidance response, rather than conditioned emotional responses which mediate avoidance learning. However, negative results where changes in heart rate do not correlate with changes in avoidance behaviour, as did occur during certain phases of this experiment, provide information which could have theoretical relevance.

Method

Subjects

The Ss were 10 adult mongrel dogs. The animals were obtained from the Harvard Medical School and had not been used in previous experimentation.

Apparatus

The apparatus (shown in Figure 1) has been described in detail in a previous paper (4). The dog rested in the hammock during avoidance learning and extinction. Its position was similar to that which it would assume when standing. In order to prevent excessive struggling, the dogs were strapped into the hammock with cloth tapes, and a wooden yoke with a semi-circular opening was secured over the neck.

The CS was a tone delivered from a Bud oscillator placed about a foot behind the hammock in which the dog rested. The US was a 60 cycle A.C. 4.5 ma. electric shock, delivered by plate electrodes attached to the dog's hind legs. Amount of current was controlled by a variable resistor in series with the animal.

The avoidance response was defined as pushing one of two metal plates which were attached to the apparatus on either side of the dog's head. The plates were hinged at the top and were free to swing out in an arc of 1.25 in. when pushed. Latency of the avoidance response was measured by a standard electric timer.

Heart rate was obtained by a modified Grass EEG polygraph. All recording and

control apparatus was kept in a room adjoining the actual experimental chamber. The experimenter observed the dog from this adjoining room through a one-way mirror. A small fan was kept going in the experimental room to mask the noise of the recording equipment.

Procedure

Acquisition. The dog was first strapped into the apparatus, and the wooden yoke attached over its neck and adjusted so that its head could move freely from side to side. *E* then left the experimental room. The dog was given a 5 min. acclimatization period followed by a series of presentations of the CS alone. These trials were given in order to extinguish orienting responses elicited by the CS, and were called *pre-*

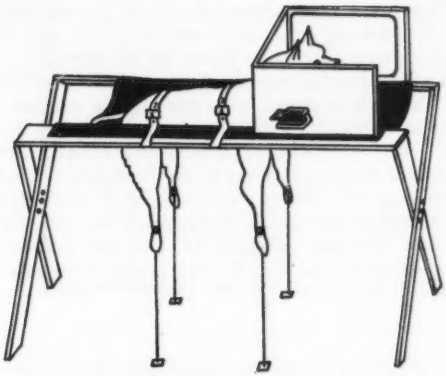

FIGURE 1. Diagrammatic sketch of the conditioning apparatus.

tests. Then avoidance training began. If the animal made the appropriate panel-pushing response within 5 sec. of CS onset, the CS was terminated and the dog avoided shock on that trial. If the animal did not make the panel-pushing response during the 5 sec. CS-US interval, the shock was turned on until the dog made the panel-pushing response, at which time both CS and shock were terminated.

The inter-trial interval averaged 1½ min. Fifteen consecutive avoidance responses were arbitrarily designated as the acquisition criterion.

Extinction. After meeting the acquisition criterion, 5 animals in the experimental group received their first 50 extinction trials while immobilized by d-tubocurarine.[2] Since they could not make the avoidance response during this period, the CS was presented for a predetermined time. The remaining 5 dogs in the control group were given 50 trials of regular extinction before curarization. They received no special treatment while under d-tubocurarine.

[2]The present paper describes observations on heart rate which were made during the course of an experiment on the effects of curarization on the extinction of an avoidance response (4). A more detailed description of the curarization procedure and the data on the instrumental avoidance response are given in this previous paper.

All 10 animals were given a 48-hour rest period following recovery from curarization. During the subsequent sessions, all dogs were extinguished in the *normal* state. They were never shocked again. The animals were run on regular extinction procedure, that is, the occurrence of the avoidance response within 5 sec. of CS onset terminated the CS. If, however, the dog did not respond, the CS was terminated by *E* after 5 sec. The inter-trial interval averaged 1 min.; intervals of 30, 60, and 90 sec. were used. Each animal was run until it failed to respond on 10 consecutive CS presentations. If this extinction criterion was not met in 400 trials, the experiment was terminated. Each animal received 100 trials a day during regular extinction.

Measurements

Heart rate was recorded on each trial for a period beginning 15 sec. before CS onset and ending 30 sec. after the occurrence of the avoidance response. If no avoidance response occurred, heart rate was recorded until 30 sec. after CS termination. These measurements were taken on all acquisition trials and on all regular extinction trials following recovery from curarization. No data are presented for the trials on which experimental and control group animals received different treatments (extinction before and during curarization) because the effect of d-tubocurarine in raising heart rate made comparisons with normal heart rate difficult.

The measure used in studying changes in the heart rate response to the CS was the comparison between the average heart rate for 5 sec. before CS onset (the pre-CS heart rate), and the average heart rate for 5 sec. after CS onset (the post-CS heart rate). The difference between the pre- and post-CS heart rates was called the *heart rate response* to the CS. The pre-CS heart rate was used as the measure of base-line rate.

Results

Acquisition

Avoidance response. A brief description of the acquisition of the avoidance response is given in Table I. The trial of first avoidance and trial of last shock are shown. The trial of last shock is assumed to be an index of speed of acquisition.

TABLE I

Selected Indices of Acquisition of the Avoidance Response

Measure	Median	Range
Trial of 1st avoidance	8.5	4–52
Trial of last shock	57.5	27–127

Base-line heart rate. The base-line heart rate showed a slight increase during acquisition (see Figures 2 and 3). The mean rate during inter-trial intervals was 113.7 beats per minute on the pre-test trials, 118.2 for acquisition trials 6 through 10, and 120.5 during the first five acquisition criterion trials. The differences in heart rate between these periods is not significant at the 5 per cent level (Method of Allowances (8)).

Heart rate responses to the CS. Since the dogs were trained until they met the acquisition criterion for the avoidance response, each animal received a different number of trials. Therefore, one can combine data only at certain reference points where it seems sensible to assume that the dogs are all roughly at the same stage of learning. Such points are the beginning of acquisition and the trials on which the acquisition criterion was met.

In Figure 2 data are shown for all dogs on the four pre-test trials (presentation of the CS alone before acquisition began) and on the first

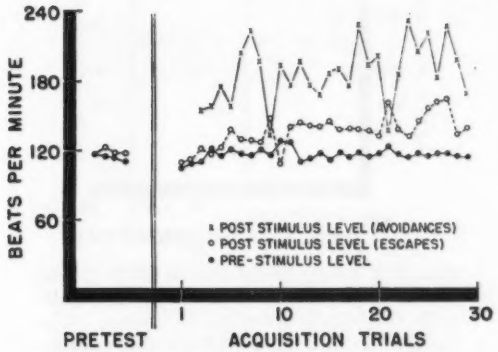

FIGURE 2. Average heart rate for all Ss during the pre-test trials and the first 30 acquisition trials. The average heart rate during the 5 sec. before CS onset and the 5 sec. after CS onset are presented for each trial. The heart rate after CS onset is plotted separately for avoidances and failures to avoid (escapes). All heart rate responses shown occur before shock onset.

thirty acquisition trials. The pre-CS heart rate and the post-CS heart rate for both avoidance and non-avoidance trials are presented. Data shown are the average heart rates for the 5 sec. periods in beats per minute. There is a gradual increase in the post-CS rate as acquisition progresses. Also the heart rate response is greater on avoidance than on non-avoidance trials.

In Figure 3 pre- and post-CS rates are presented for the fifteen acquisition criterion trials and the ten trials immediately preceding. The heart rate response to the CS is much greater on trials following the trial of last shock, when the avoidance response has been learned, than on the pre-test trials. The mean heart rate response on the first five criterion

trials is 33.4 beats per minute, and on the pre-tests 3.6 beats per minute. This difference is significant at the 5 per cent level (Wilcoxon's signed rank test for pair replicates).

The measure of heart rate responding used above (average heart rate during a 5 sec. period) by no means reflects moment-to-moment changes

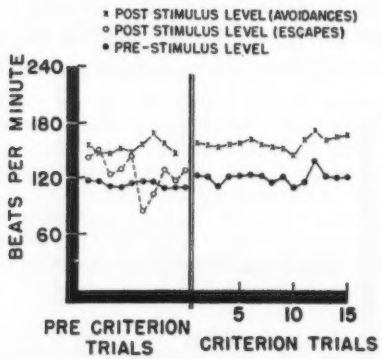

FIGURE 3. Average heart rate for all Ss during the 15 acquisition criterion trials and the 10 trials immediately preceding the criterion trials. See the legend of Figure 2 for further explanation of these data.

in cardiac behaviour. In Figure 4 heart rates for a typical trial are presented. The instantaneous rate[3] (in beats per minute) is plotted for 10 sec. before CS onset and 10 sec. after CS onset. The solid horizontal bar in the centre of the graph represents CS duration. When the CS is presented, there is a sharp acceleration in rate which reaches its maximum 2 to 3 sec. *after* the CS is terminated by the avoidance response and this is followed by a sudden drop in rate. The beat-to-beat variability of the heart rate is greater at this time than it was just prior to CS presentation.

An analysis of grouped data for the last five acquisition criterion trials showed that the heart rate response did not reach its maximum until 1.5 sec. after the avoidance response occurred, and did not begin to decline until 2.8 sec. after the occurrence of the avoidance response.

During acquisition, the most common behaviour was that pictured in the example in Figure 4. In addition to this pattern of responding, there

[3]Instanteous rate is found by multiplying the reciprocal of the time between each QRS complex by 60. For example, if the time between two QRS complexes is .50 sec., the instantaneous rate would be $\frac{1}{.50} \times 60 = 120$ beats per minute.

were trials in which *no* change was apparent when the CS was presented, and some trials on which a decrease was noted. In only one dog did this deceleration to the CS occur frequently.

Heart rate response and general activity. There was a close correlation between magnitude of heart rate response and amount of skeletal activity

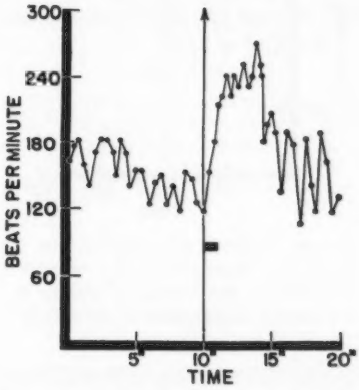

FIGURE 4. Beat-to-beat changes in heart rate on a single trial during the latter stages of acquisition. The solid horizontal bar in the centre of the figure represents Cs duration. CS onset is represented by the vertical line.

such as struggling and barking. While no direct measures of activity were made, notes of behaviour were taken on each trial and activity was ranked on a four-point scale from no activity through slight and medium activity to violent activity. The effect of violent struggling was to increase the magnitude and the duration of the heart rate response. An example of a trial on which such violent activity occurred is shown in Figure 5. The magnitude and duration of the response on this trial are much greater than those on the trial shown in Figure 4 where little extraneous activity was observed.

In describing Figure 2, we noted that heart rate responses were higher on avoidance than on non-avoidance trials. If data from these trials are broken down into three categories, avoidance trials, non-avoidance low activity trials, and non-avoidance high activity trials, it becomes clear that high heart rate is correlated with activity as well as avoidance responses *per se*. The data in Table II show heart rate responses for each of these categories on the first three trials following the first avoidance.

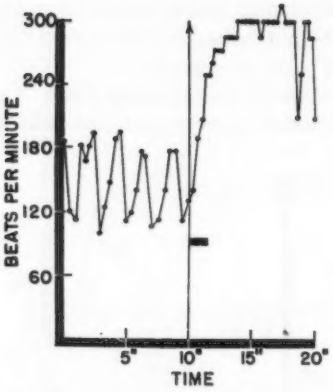

FIGURE 5. Heart rate responses on a single trial during the latter stages of acquisition. See legend of Figure 4 for further explanation of these data.

There is no significant difference in heart rate between the avoidance and non-avoidance high activity trials, but we do recognize a difference between non-avoidance low activity trials and trials in the other two categories at the 5 per cent level (Method of Allowances).

TABLE II

HEART RATE RESPONSE DURING ACQUISITION (IN BEATS PER MINUTE)

	Non-avoidance trials	
Avoidance trials	High activity	Low activity
36.2	37.4	15.1

There was considerable trial-to-trial variability in the magnitude of the heart rate response during acquisition and this variability in heart rate was associated with variability in general activity. During the early stages of acquisition, activity alternated in a seemingly random pattern with periods of inhibition of movement. Such inhibition was an active process rather than a failure of movement since attempts to move the dog's head or legs during this period were met with great resistance. This pattern of behaviour is similar to that found by Gibson (6) in her observations on the classical conditioning of skeletal responses in goats. During extinction the amount of activity decreased rapidly to a low level.

Extinction

The avoidance response. Half of the animals (experimental group) received fifty extinction trials while curarized, and half received fifty regular extinction trials and were then curarized (control group). Following these procedures, all the dogs were run on regular extinction. The data on regular extinction following this differential treatment are shown in Table III. Dogs in the experimental group reached the extinction criterion

TABLE III

INDICES OF EXTINCTION OF THE AVOIDANCE RESPONSE

Measure	Experimental group		Control group	
	Median	Range	Median	Range
No. of trials of regular extinction	41	7–79	450+	98–450+
Responses in extinction	24	6–47	449	70–450
Avoidances on first 15 extinction trials	7	4–15	15	12–15
Latency of response on first 15 extinction trials	5.0+	0.4–5.0	1.1	0.4–2.0

more rapidly and made fewer avoidances during regular extinction than dogs in the control group. This difference between groups is apparent on the early extinction trials since they differ significantly at the 5 per cent level on number of responses and latency of response on the first fifteen extinction trials.

Base-line heart rate. Although there were clear-cut differences in avoidance responding between experimental and control groups on the regular extinction trials, no such differences were found for base-line heart rate. Table IVA shows the mean base-line heart rate for the two groups on the fifteen criterion trials and on the first fifteen extinction trials following the differential treatment. If this treatment had an effect on base-line heart rate similar to that on the avoidance response we would expect a significant interaction between group and block of trials. That is, there should be a more pronounced difference in heart rate between criterion and extinction trials among experimental group dogs than control group dogs. However, as can be seen from the analysis of variance shown in Table IVB, this interaction is not significant.

During the first extinction session there was a rapid drop in the base-line heart rate. The mean base-line rate at the beginning of the first extinction session was 124.1 beats per minute, and the rate at the end was 97.7. This difference is significant at the 5 per cent level (Wilcoxon's signed rank test for paired replicates). Base-line rates showed spontaneous recovery at the beginning of each extinction session. These changes are

TABLE IV

A. A Comparison of Base-line Heart Rate Response at the End of Acquisition and Beginning of Regular Extinction

	Base-line heart rate* during 15 criterion trials	Base-line heart rate* at beginning of regular extinction
Experimental group	126.8	120.1
Control group	116.1	117.4

*Rates are in beats per minute.

B. Analysis of Variance of Base-line Heart Rate Data

Source	S.S.	df	M.S.	F
Total	949.20	19		
Blocks of trials	2.21	1	2.21	$F < 1$
Groups	13.86	1	13.86	$F < 1$
Within groups	840.14	8	105.02	
Group × blocks of trials	5.05	1	5.05	$F < 1$
Block of trials × dog within group	87.94	8	10.99	

illustrated in Figure 6 where pre- and post-CS heart rates during extinction for a single dog are shown.

Heart rate response to the CS. A similar analysis of variance was carried out in order to determine whether differential treatment had any effect on the heart rate response to the CS. The data used in this analysis are

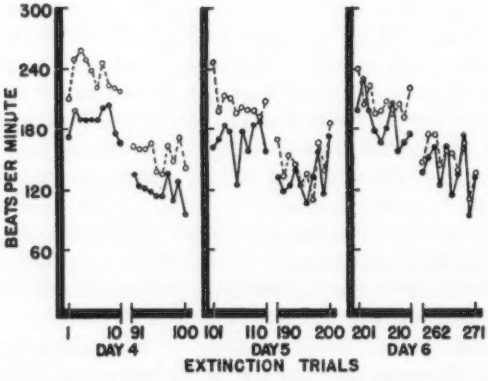

Figure 6. Heart rate for a single S during extinction trials. The solid dots represent the average rate during the 5 sec. before CS onset on each trial. The open circles represent the average heart rate during the first 5 sec. after CS onset on each trial.

shown in Table V. Again the interaction between group and block of trials is not significant. Extinction trials under curare did not seem to have the same differential effect on the heart rate response as they did on the avoidance response.

TABLE V

A. A Comparison of Heart Rate Response to the CS at the End of Acquisition and Beginning of Regular Extinction

	Heart rate* response during 15 criterion trials	Heart rate* response at beginning of regular extinction
Experimental group	38.4	16.08
Control group	38.88	18.48

*Rates are in beats per minute.

B. Analysis of Variance of Heart Rate Response Data

Source	S.S.	df	M.S.	F
Total	1878.6474	19		
Blocks of trials	387.2880	1	387.2880	$Pr(F=10.97)<.05$
Groups	2.0162	1	2.0162	$F<1$
Within groups	1205.6670	8	150.7084	
Group × Blocks of trials	1.3364	1	1.3364	$F<1$
Block of trials × dog within group	282.3398	8	35.2925	

In all cases, the heart rate response extinguished before the avoidance response. For example, by defining a heart rate response arbitrarily as one greater than the average pre-test response, the number of trials required to reach an extinction criterion could be obtained. When five failures to respond were noted in any block of ten trials, extinction of the cardiac response was said to have occurred. The median number of trials to reach extinction criterion for the *cardiac* response was ten (range 0–245); the median number of trials to reach the extinction criterion for the *avoidance* response was 63.5 (range 7–450).

Since the heart rate response extinguished before the avoidance response, it is important to determine whether the speed of extinction of the heart rate response could be used as a predictor of speed of extinction of the avoidance response. Therefore, a number of different criteria for the extinction of the cardiac CR, in addition to that used above, were employed in the hope that one might yield a significant result which could be verified in further studies. However, the highest correlation found between speed of extinction of the avoidance and cardiac CRs was only + .41, and not significant.

Discussion

The close correlation between skeletal movement and magnitude of the heart rate response during the acquisition phase of this experiment suggests that Peters and Gantt (**10**) were incorrect in minimizing the possibilities for contamination of heart rate data by skeletal responding.

One could argue, however, that this whole problem is not really serious. We could simply assume that skeletal behaviour along with heart rate and many other responses form a patterned complex which we label "emotional" responding as Solomon and Wynne have done (**13**). If the interrelated responses in this syndrome are highly correlated with each other, this is only to be expected.

That such a pattern of autonomic and skeletal responses exists is in all probability correct. However, in studying conditioning it is also probable that different responses in this syndrome are sensitive to control by different types of conditioning variables. This is particularly true of skeletal behaviour and heart rate. For example, the heart rate response showed great trial-to-trial variability during acquisition. This does not seem to be the typical pattern for the development of conditioned cardiac responses. When cardiac conditioning is carried out in dogs whose skeletal musculature has been paralysed by curare-like drugs (**3**), there is a gradual increase in magnitude of response from trial to trial until an asymptote is reached, and the trial-to-trial variability is extremely low.

This discrepancy indicates that the heart rate variability in the present experiment could be attributed to the effect of variables which influence skeletal activity rather than heart rate directly. In this particular case, it seems plausible to suggest that skeletal components of the emotional response were punished when they occurred immediately prior to shock onset, while heart rate responses were not affected by the US as a punisher. This would lead to periods of alternating activity and inhibition which would in turn lead to variability in heart rate.

Similarly the avoidance response itself, which is one form of skeletal activity, is controlled by reinforcement variables such as CS termination, which we normally do not manipulate in order to control conditioned emotional responses.

Thus, although both skeletal responding and heart rate originally are part of the same set of emotional responses, it is possible that they are influenced directly by different types of variables. If variables which control skeletal behaviour, but not heart rate, are operating in a given learning situation, as is probably the case in avoidance learning, the interpretation of results remains ambiguous. Therefore, we cannot decide whether changes in heart rate are controlled directly or indirectly by stimulus events during learning.

This is not to suggest that all heart rate changes observed in this experiment were simply artifacts of activity. While there was a close correlation between heart rate and activity in acquisition, during the extinction phase of the experiment there were clear differences between groups in avoidance responding but no such differences in the heart rate response.

If we use heart rate as an index of the autonomic responses which Mowrer discusses in his theory, then these results indicate that the maintenance of avoidance responding during extinction does not depend on the maintenance of conditioned emotional responses to the CS, as he has postulated. However, it could be argued that our failure to find positive results during extinction could be attributed to the high variability of heart rate behaviour found in this particular experiment, and that another less "noisy" measure of autonomic behaviour or a different experimental situation would have been more suitable. But even if such positive results were found as was the case in the acquisition phase of this experiment, and in the work of Bersh, Schoenfeld, and Notterman on human subjects (2), the criticisms outlined in the earlier part of this discussion are relevant. Thus it seems clear that theories such as Mowrer's, which simply treat autonomic behaviour as an index of fear or anxiety without regard to the central and peripheral interrelations between autonomic and skeletal behaviour are too simple to explain our findings.

Summary

Ten animals were first trained to avoid an intense shock, and then all *Ss* were run until the avoidance response extinguished. Heart rate responses were measured on each trial. The heart rate response to the CS was typically an increase in heart rate. There was a relationship between the intensity of the heart rate response and the occurrence of the avoidance response during acquisition. Also, skeletal activity and heart rate responding were correlated during acquisition. There was no significant correlation between speed of extinction of the cardiac and avoidance responses.

The implications of these results for the measurement of conditioned emotional responding were discussed.

REFERENCES

1. Ax, A. The physiological differentiation between fear and anger in humans. *Psychosom. Med.*, 1953, **15**, 433–442.
2. Bersh, P. J., Notterman, J. M., & Schoenfeld, W. N. Extinction of a human cardiac response during avoidance conditioning. *Amer. J. Psychol.*, 1956, **69**, 244–251.
3. Black, A. H., & Carlson, N. J. The conditioning of autonomic responses under curare-like drugs. *Amer. Psychol.*, 1957, **12**, 429.

4. BLACK, A. H. The extinction of avoidance responses under curare. *J. comp. physiol. Psychol.*, 1958, **51**, 519–524.
5. GANTT, W. H. Principles of nervous breakdown-schizokinesis and autokinesis. *Ann. N.Y. Acad. Sci.*, 1953, **56**, 143–163.
6. GIBSON, E. J. The role of shock in reinforcement. *J. comp. physiol. Psychol.*, 1952, **45**, 18–30.
7. MCLEARY, R. A. Measurement of experimental anxiety in the rat: An attempt. *J. genet. Psychol.*, 1954, **84**, 95–108.
8. MOSTELLER, F., & BUSH, R. R. Selected quantitative techniques. In G. LINDZEY (Ed.), *Handbook of social psychology*, Cambridge: Addison-Wesley, 1954.
9. MOWRER, O. H. On the dual nature of learning—a reinterpretation of "conditioning" and "problem-solving", *Harv. educ. Rev.*, Spring, 1947, 102–148.
10. PETERS, J. E., & GANTT, W. H. Effect of graded degrees of muscular exertion on human heart rate and the role of muscular exertion in cardiac conditional reflexes. *J. gen. Psychol.*, 1953, **49**, 31–43.
11. SMITH, K. Conditioning as an artifact. *Psychol. Rev.*, 1954, **61**, 217–225.
12. SOLOMON, R. L., & WYNNE, L. C. Traumatic avoidance learning: acquisition in normal dogs. *Psychol. Mongr.*, 1953, **67**, 1–19.
13. SOLOMON, R. L., & WYNNE, L. C. Traumatic avoidance learning: the principle of anxiety conservation and partial irreversibility. *Psychol. Rev.*, 1954, **61**, 353–385.

THE EFFECTS OF DIFFERENTIAL INFANTILE STIMULATION ON EMOTIONALITY AT WEANING[1]

SEYMOUR LEVINE

Columbus Psychiatric Institute and Hospital
Ohio State University

DURING THE PAST several years there has been a growing literature related to the problem of the effects of early experience on later behaviour. Investigators have reported improved learning, reduced emotionality and resistance to stress as a function of experience during the pre-weaning and immediately post-weaning periods.

Various types of treatment have been administered. Among them are: electric shock, Levine, Chevalier, and Korchin (6); gentling, Weininger (15); manipulation, Levine (5); stroking, McClelland (11); and partial starvation, Hunt (4). Although the nature of infantile experience has been seemingly diverse, there is one common element present in all the treatments, namely that it has been necessary to handle the animal in order to initiate the experimental treatments. Until very recently all the studies dealing with the effects of infantile experience have utilized procedures which involve some contact with the experimenter.

However, a recent study by Levine and Lewis (9) showed that infant animals which were either shaken or moved about during the first thirteen days of life (both procedures not involving handling) revealed significantly greater depletion of adrenal ascorbic acid at fourteen days of age whereas the non-treated control animals did not show depletion of ascorbic acid under identical test conditions.

The present experiment was designed to determine (1) the effects of infantile stimulation in the absence of handling on emotionality as measured by defecation and urination in a novel situation, and (2) whether differences obtained as a function of infantile stimulation can be detected at the time of weaning, the time when many studies dealing with the variable of early experience initiate treatment.

In addition to the above considerations the present experiment investigated the possible effects of removing the mothers under conditions of stimulation. In the previous experiment (9), although the stimulated

[1]This investigation was supported in part by research grant PHS M–1630 from the National Institute of Mental Health of the National Institutes of Health, U.S. Public Health Service.

infants were not handled, the mothers were removed from the cages daily. This procedure was not followed for the non-stimulated groups. In order to assess the possible effects of removing the mother, one group was stimulated while leaving the mother in the cage, and another group was equivalent to the non-handled group with the exception that the mother was removed daily.

Method

Subjects

As subjects, 141 infant Sprague Dawley albino rats were used. The Ss came from 15 females bred in this laboratory. They were assigned to one of 5 conditions as they were born. All Ss were born and housed in $9 \times 15 \times 9$ in. cages. In order to maintain a relatively constant home cage environment, the cages were covered with a galvanized iron top so that all sides of the cage, with the exception of the front, were enclosed.

Procedure

The 5 conditions in this experiment were as follows:

1. Shaker ($N = 30$). Once daily, starting one day after birth, the whole cage was removed from the rack on which the cage was housed. The cage cover, water bottle, and mother were removed. The cage was then placed on an Eberbach Laboratory Shaker and shaken for 2 min. at 180 oscillations per min. The cage cover, water bottle, and mother were then replaced and cage returned to the rack. At no time until the test day were the Ss handled. The mother was removed only after she had left the infants. At no time were the infants removed from the nipples by hand.
2. Shaker Control (M) ($N = 26$). This group received the same treatment as the Shaker group, but were placed on the shaker for two min. without shaking.
3. Shaker Control ($N = 29$). This group received identical treatment to the above group, with the exception that the mother was never removed from the cage during stimulation.
4. Control (M) ($N = 25$). These Ss received no manipulation. However, the mothers were removed daily and remained out of the cage for 4 min.
5. Control ($N = 31$). The control Ss remained in the breeding cages and received no treatment until the test day.

On day 21 all of the Ss were placed in a $3 \times 4 \times 6$ in. compartment for 3 min. During this period the incidence of defecation was recorded for all groups.

Results

Table I presents the number of animals in each group which defecated during the test period. A 2×5 χ^2 yielded a value of 20.76 which is significant beyond the .001 level. The difference in incidence of defecation between the Shaker, Shaker Control, and Shaker Control (M) Ss was not significant. However, the non-stimulated Ss showed a significantly greater incidence of defecation ($P < .001$) when placed in a novel situation. The Control and Control (M) Ss did not differ significantly from each other.

In addition to the above observations, it was noticed that the stimulated rats showed a marked acceleration of general physical development. Almost all the stimulated subjects had their eyes open by the fourteenth day, and in the Shaker group the opening of the eyes was observed as early as twelve days. The non-stimulated infants did not show opening of the eyes until about the sixteenth day. The same differential develop-

TABLE I

DEFECATION DURING TEST PERIOD

Group	*N*	No. defecating	Percentage
Shaker	30	9	30
Shaker Control (M)*	26	10	38
Shaker Control	29	11	38
Control (M)	25	20	80
Control	31	21	68

$X^2 = 20.76$; $P = < .001$. *(M) refers to removal of the mother.

ment was observed with respect to other aspects of development, such as appearance of fur and general locomotion. Since this study precluded handling the Ss in any manner these observations could not be quantified at this time. However, they are in accord with previous studies (**7**), (**10**) that have indicated more rapid maturation of physiological characteristics as a result of infantile stimulation.

DISCUSSION

It is apparent from the above results that handling *per se* is not the critical factor which leads to the reduced emotionality and accelerated development observed as a consequence of infantile experience.

In this study as in several previous studies, no differences in behaviour were observed between the different modes of stimulation. Although some differences have been noted in adulthood between animals that were shocked in infancy and those that were handled (**6**), these differences have been transitory, and the magnitude of the differences have been much smaller than the differences between stimulated and non-stimulated animals. The evidence obtained thus far has raised the question of what in this area constitutes the major experimental treatment. It appears that the condition of no treatment seems to have the most profound effect upon development and the subsequent emotionality and performance of the animal. Thus far in all of the experiments coming out of our laboratory it has been the non-stimulated infant that has exhibited relatively

slower development, greater emotionality, and poorer performance in adulthood.

Thus it appears that one of the main contributions of research in the area of infantile experience has been to demonstrate empirically the importance of stimulation during critical periods in development on many subsequent psychophysiological factors. The importance of infantile stimulation has, of course, been emphasized by Hebb (3). There are, however, still many problems related to the nature of the stimulation and the physiological mechanisms of the action of infantile stimulation. Both Hebb and Bovard (1) have offered physiological hypotheses for the effects of infantile experience.

The results of this study indicate further that the removal of the mother had no effect on the phenomena studied. The data support those obtained by Shaefer (14). Shaefer found that separation from the mother did not contribute to any change in emotionality in later life. Thus the effects of infantile stimulation on later emotionality appears to be a direct effect upon the developing organism and not mediated via the mother.

These data, taken in conjunction with other data related to critical periods (2, 8, 13) tend to emphasize again the point made in several previous papers (6, 7, 12), namely, that in order to evaluate the effects of early experiences it is essential that the whole life history of the organism be known. Thus the studies which use as their subjects weanling rats usually obtained from and shipped by commercial breeders add an additional degree of contamination and confusion to an area long beset with enough difficult problems.

Summary

Three groups of newborn albino rats were stimulated by being shaken or moved about for 20 days. These procedures involved no handling. At 21 days of age these Ss were compared with non-stimulated control groups for differences in emotionality.

The results revealed that the non-stimulated Ss exhibited significantly greater defecation in a novel situation than did either of the stimulated groups. In addition, the removal of the mother proved to have no effect on later emotionality.

These results were interpreted as indicating the importance of stimulation in the absence of handling on physiological development and on development of differential emotional responsivity.

REFERENCES

1. Bovard, E. W. The effects of early handling on viability of the albino rat. *Psychol. Rev.*, 1958, 65, 257–271.
2. Denenberg, V. Effects of critical periods and the amount of early experience on adult learning. Paper delivered at the American Psychological Association, Washington, D.C., September 2, 1958.

3. HEBB, D. O. *The organization of behavior.* New York: Wiley, 1949.
4. HUNT, J. McV. The effects of infant feeding frustration in the albino rat. *J. abnorm. soc. Psychol.*, 1941, **36**, 338–360.
5. LEVINE, S. A further study of infantile handling and adult avoidance learning. *J. Pers.*, 1956, **25**, 70–80.
6. LEVINE, S., CHEVALIER, J. A., & KORCHIN, S. J. The effects of shock and handling in infancy on later avoidance learning. *J. Pers.*, 1956, **24**, 475–493.
7. LEVINE, S., ALPERT, M., & LEWIS, G. W. Differential maturation of an adrenal response to cold stress in rats manipulated in infancy. *J. comp. physiol. Psychol.*, 1958, **51**, 774–777.
8. LEVINE, S., & LEWIS, G. W. Critical period for effects of infantile experience on maturation of stress response. *Science*, 1959, **129**, 42–43.
9. LEVINE, S., & LEWIS, G. W. The relative importance of experimenter contact in an effect produced by extra stimulation in infancy. *J. comp. physiol. Psychol.*, 1959 (In press).
10. LEVINE, S., & ALPERT, M. Differential maturation of the central nervous system as a function of prior experience. *Arch. Gen. Psychiat.*, 1959, **1**, 403–405.
11. MCCLELLAND, W. J. Differential handling and weight gain in the albino rat. *Canad. J. Psychol.*, 1956, **10**, 19–22.
12. ROSEN, J. Comments on "Studies of the effects of early gentling experience." *Psych. Rep.*, 1958, **4**, 623–626.
13. SCOTT, J. P. Critical periods in the development of social behavior in puppies. *Psychosom. Med.*, 1957, **20**, 42–53.
14. SHAEFER, T. The effect of mother, infant separation on adult behavior. Annual Report. Institute for Psychosomatic and Psychiatric Research and Training, Michael Reese Hospital, Chicago, 1957, **12**, 54 (Abstract).
15. WEININGER, O. The effects of early experience on behavior and growth characteristics. *J. comp. physiol. Psychol.*, 1956, **49**, 1–9.

EFFECTS OF PAST EXPERIENCE ON EXPLORATORY BEHAVIOUR IN RATS[1]

ANNETTE EHRLICH

McGill University

EXPLANATIONS of exploratory behaviour in rats frequently involve the postulation of an exploratory drive, the distinguishing feature of which is that it is a response to exteroceptive stimulation only. Montgomery, for example, devoted a whole series of experiments to demonstrating that exploratory activity not only is independent of homeostatic drives such as hunger and thirst (6) but also has little or nothing to do with the past experience of the animal (7, 9). Yet, after reviewing the literature on this topic, it seemed to the present writer that the issue of whether or not past experience has any effect on exploratory behaviour had not been fully resolved. For one thing, the experiments dealing with early rearing environments left much to be desired since none of them involved a direct comparison between free-environment and restricted rats, but involved instead a comparison between an experimentally reared group (free or restricted) and a normally reared control group (4, 9, 11). The data relating to the effects of immediate past experience were similarly equivocal, for despite Montgomery's claim that exploratory activity shows "spontaneous recovery" between daily trials, his data are actually far from uniform,[2] and they conflict with the results of several recent experiments (3, 11, 9[3]). It seemed worthwhile, therefore, to evaluate the effects on exploration of (*a*) early environment and (*b*) immediate past experience. Specifically, the effects on exploration of free *versus* restricted rearing environments, handling *versus* non-handling, and of repeated exposures to the same test situation were investigated. In addition, it was decided to use two mazes, a Y and a Dashiell, to test a suggestion advanced previously by Zimbardo and Montgomery (11) that the effects of differential rearing conditions may

[1]This research was supported (in part) by grant no. M-2455 from the U.S. Public Health Service, National Institute of Mental Health, to Dr. D. O. Hebb. The author is indebted to Mr. D. Ehrlich for designing and constructing the test apparatus.

[2]In 1953, for example, in a single paper (5), he reported a between-trials increase in experiment I and a between-trials decrease in experiment II.

[3]The data for the between-days decline in this study are reported in Zimbardo and Montgomery. (11).

vary with the complexity of the test situation. These authors pointed out that the Y maze may not be sufficiently novel to elicit a high degree of exploration by animals raised in an enriched environment, and they suggested the use of a more complex maze with these animals. In line with their analysis, it was predicted that the free-environment animals would explore less than the restricted in the Y maze but more than the restricted in the Dashiell maze.

Method

Subjects

The subjects were 40 male hooded rats obtained at the time of weaning from the Royal Victoria Hospital colony. The animals were maintained on an *ad lib.* diet of Purina Fox Chow pellets and water throughout the experiment.

Apparatus

Rearing. The 20 free-environment animals were housed in a two-tiered cage which measured 30 × 48 × 24 in.; each level was 12 in. in height. There were various playthings such as corks, tunnels, platforms, and inclined ramps in the cage. The walls were made of wire mesh so that the animals could see the rest of the room. The 10 handled animals occupied one level and the 10 non-handled animals occupied the other level, the location of the subgroups being interchanged weekly. The 20 restricted rats (half of whom were handled) were reared in 4 colony cages measuring 24 × 12 × 9 in., the doors and tops of which were covered with white paper to prevent the animals from seeing out. There were 5 rats in each cage, but no playthings were provided.

Testing. The Y maze used was similar to that described by Montgomery (6). Each arm measured 24 × 4 × 4 in. and was marked off into 12 in. units. The maze arms were constructed of ¾ in. unpainted pine and were covered with hardware cloth. The Dashiell maze (2) was an enclosed maze arranged in a checkerboard pattern; it measured 28 × 28 × 4 in. The alleys were 4 in. wide and were divided into 4 in. squares. The walls and floors were constructed of ½ in. unpainted pine and the top was of hardware cloth. Lighting for both mazes was provided by a single 100-watt bulb suspended directly over the centre of the maze.

Procedure

Rearing. At the age of 21 days, the 40 animals were randomly divided into 4 groups—free-environment handled, free-environment non-handled, restricted handled, and restricted non-handled. Handled animals were removed from their cages once every day and stroked 10 times. Non-handled animals were picked up briefly once a week when their cages were cleaned. Since the object was to measure the effect of differential rearing conditions rather than the effect of differential caging conditions immediately prior to exploration, all animals were transferred to a uniform environment a week before the start of formal testing and were housed two to a cage in standard laboratory cages measuring 10 × 12 × 9 in. The animals could see part of the room through the wire mesh doors of these cages. During the 3 days prior to the start of testing, each animal was handled for 2 min. per day. The animals were 76 days old at the start of testing.

Testing. All animals were run on both mazes; and, in order to control for the possible effect which exploration of one maze might have on exploration of the other, the 40 animals were divided into 2 groups of 20 each. Group I contained 5 animals from each of the 4 subgroups. Each animal in this group was tested individually in the Y maze for 10 min. per day on 3 successive days and in the Dashiell maze for 10 min. per day on the succeeding 3 days. Group II, which also consisted of 20 animals with 5 from each subgroup, was tested 6 days after group I; for this group, the procedure was the same except that experience with the Dashiell maze preceded testing in the Y maze. The measure of exploration in each maze was the number of units entered during each of five 2 min. intervals. The placing of head and forepaws within a unit constituted the criterion for entry into a maze unit. In the Y maze, all rats were started at the end of one arm; in the Dashiell maze, all rats were started at a central point on one side of the maze.

Results

Total Exploratory Activity

Tables I and II show the analyses of variance for the scores in the Y and Dashiell mazes respectively for the three trials combined. In both mazes, handling was the only variable which significantly affected exploration. Handled animals explored more than non-handled animals

TABLE I

Analysis of Variance of Amount of Exploratory Behaviour in Y Maze for All Three Trials Combined

Source	df	Mean square	*F*	*P*
Between rearing (R)	1	5499.1	2.02	
Between handling (H)	1	18792.3	6.91	< .05
Between maze experience (E)	1	2030.7	.75	
R × H	1	4.2	.001	
R × E	1	3115.2	1.15	
H × E	1	1311.0	.48	
R × H × E	1	189.1	.06	
Within groups	32	2718.3		

TABLE II

Analysis of Variance of Amount of Exploratory Behaviour in Dashiell Maze for All Three Trials Combined

Source	df	Mean square	*F*	*P*
Between rearing (R)	1	36240.5	1.20	
Between handling (H)	1	212868.1	7.04	< .05
Between maze experience (E)	1	656.1	.02	
R × H	1	7617.6	.25	
R × E	1	15523.0	.51	
H × E	1	9180.9	.30	
R × H × E	1	9486.3	.31	
Within groups	32	30243.3		

($P < .05$ in both cases). Also, there was a tendency for the restricted animals to explore more than the free-environment animals although the F ratio for rearing conditions did not reach the .05 level of significance for the scores in either maze. Prior experience in one maze had no effect on performance in the other maze. Maze complexity did not affect exploration. Not only were the group trends in the Y and Dashiell maze data the same, but also the rank-order correlation coefficient between performance in the two mazes was .88 (which differed significantly from zero at the .01 level).

Between-days Decline

Figures 1 and 2 show the mean number of maze units traversed by handled and non-handled animals for the three test trials in the Y and Dashiell mazes respectively. It is apparent from these figures that

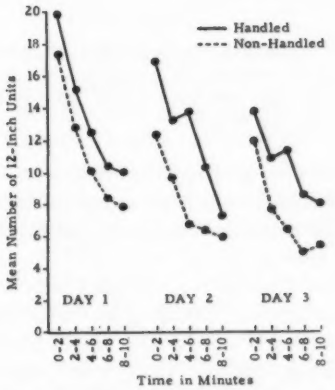

FIGURE 1. Mean number of 12 in. units entered in Y maze by handled and non-handled rats as a function of time.

exploration declined both within and between trials. For the group as a whole, significant declines were noted between the first and third trials in both the Y and the Dashiell maze (t-test for correlated means, $P < .001$ in both cases). However, the amount of decline was not uniform among the various subgroups. Analyses of variance of the difference scores between the first and third trials on each maze yielded the following results.

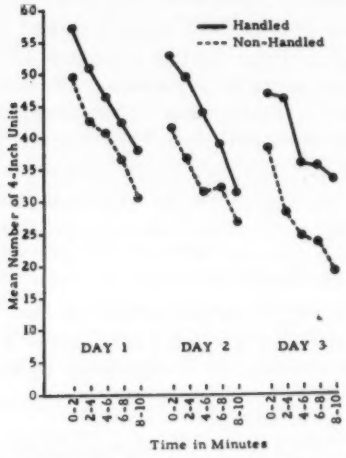

FIGURE 2. Mean number of 4 in. units entered in Dashiell maze by handled and non-handled rats as a function of time.

Y maze. The F ratio for rearing conditions was 8.40 ($P < .01$). The exploratory scores of free-environment animals declined less rapidly than those of the restricted animals. Handling and maze experience did not affect the between-days decline.

Dashiell maze. The F ratio for handling was 5.25 ($P < .05$). The significant interaction between handling and maze experience ($F = 4.68$, $P < .05$) showed that this was primarily due to the fact that the exploration of handled animals without prior maze experience declined less sharply than that of the other groups.

DISCUSSION

The results indicate that handling significantly increases exploration, a finding which is consistent both with Bernstein's incidental observation that handled rats explore more than non-handled in a maze-learning situation (1) and with Montgomery and Monkman's findings on the inverse relationship between fear and exploration (8). These data also suggest that the greater degree of exploration shown by the free-environment animals in an experiment reported by Luchins and Forgus (4) might have been due to the fact that their experimental animals were handled whereas their normal controls were not.

The effects of the free and restricted environments are, unfortunately, less clear-cut. Owing to the small size of the groups and to the large degree of variability, the *F* ratios for rearing conditions failed to reach an acceptable level of significance. Yet, the tendency for restricted rats to explore more than free-environment rats was quite consistent, since it occurred in both mazes and in groups with and without prior maze experience. Furthermore, in both mazes, the group with the lowest exploration score was the free-environment non-handled group. Although the present results do not demonstrate conclusively that free and restricted rearing environments have differential effects on exploratory behaviour, there is some support for this position in a recent experiment by Woods (**10**). He compared the exploratory behaviour of free-environment and restricted rats in a Hebb-Williams maze, taking as his measure of exploration the number of times an animal retraced its path after reaching the goal box. Woods' results, like those of the present experiment, indicated that restricted rats explore more than do rats raised in a free environment.

Contrary to expectations, the data did not show that maze complexity was a factor in the amount of exploratory activity shown by free-environment and restricted rats. However, it is possible that the Dashiell maze was not sufficiently complex. A maze with several dissimilar alternatives might have elicited more exploration by the free-environment animals.

With respect to changes in exploratory behaviour as a function of repeated trials in the same maze, it was found that all animals explored significantly less on the third than on the first trials in both mazes. Although this finding is inconsistent with Montgomery's early reports (**5, 6, 7**), it is in agreement with several recent experiments (**3, 9, 11**), and it seems reasonable to conclude that exploration in the Y maze by satiated rats tends to decay both within and between trials. It should be noted that the rate of between-days decline was not the same in all groups, but the meaning of this finding is not clear. Handling resulted in significantly less decline in the Dashiell maze and free-environment rearing resulted in significantly less decline in the Y maze. It is interesting to note that, despite the significant decline with repeated exposures to the same maze, prior experience in one maze had no effect on exploration in the other maze.

Summary and Conclusions

The effects on exploration of differential rearing conditions and of immediate past experience were studied and, in addition, the factor of maze complexity was varied by utilizing two tests, a Y maze and a Dashiell maze. Early handling significantly increased exploration in both mazes, and there was a tendency, although

not significant, for rats reared in a restricted environment to explore more than rats reared in a free environment. Maze complexity did not affect the results. Although previous experience in one maze did not affect performance in the other maze, exploratory activity declined with time, both within and between trials in the same maze. The results justify the statement that a rat's response to a novel situation is dependent upon its past history, both in terms of early experience and in terms of previous exposures to the same situation.

REFERENCES

1. Bernstein, L. A note on Christie's "Experimental naïveté and experiential naïveté." *Psychol. Bull.*, 1952, **49**, 38–40.
2. Dashiell, J. F. A quantitative demonstration of animal drive. *J. comp. Psychol.*, 1925, **5**, 205–208.
3. Glickman, S. E. Effects of peripheral blindness on exploratory behaviour in the hooded rat. *Canad. J. Psychol.*, 1958, **12**, 45–51.
4. Luchins, A. S., & Forgus, R. H. The effect of differential post-weaning environments on the rigidity of an animal's behavior. *J. genet. Psychol.*, 1955, **86**, 51–58.
5. Montgomery, K. C. Exploratory behavior and its relation to spontaneous alternation in a series of maze exposures. *J. comp. physiol. Psychol.*, 1952, **45**, 50–57.
6. Montgomery, K. C. The effect of the hunger and thirst drives upon exploratory behavior. *J. comp. physiol. Psychol.*, 1953, **46**, 315–319.
7. Montgomery, K. C. The effect of activity deprivation upon exploratory behavior. *J. comp. physiol. Psychol.*, 1953, **46**, 438–441.
8. Montgomery, K. C. & Monkman, J. A. The relation between fear and exploratory behavior. *J. comp. physiol. Psychol.*, 1955, **48**, 132–136.
9. Montgomery, K. C., & Zimbardo, P. G. Effect of sensory and behavioral deprivation upon exploratory behavior in the rat. *Percept. Mot. Skills*, 1957, **7**, 223–229.
10. Woods, P. J. The effects of free and restricted environmental experience on problem solving behavior in the rat. *J. comp. physiol. Psychol.*, 1959, **52**, 399–402.
11. Zimbardo, P. G., & Montgomery, K. C. Effects of "free environment" rearing upon exploratory behavior. *Psychol. Rep.*, 1957, **3**, 589–594.

INTERACTION BETWEEN PROACTION AND RETROACTION IN SHORT-TERM RETENTION[1]

ENDEL TULVING AND GERALD B. THORNTON

University of Toronto

WHEN A PERSON learns something and some time later is given a retention test on this material, his performance can be shown to depend on a large number of variables. Two important classes of these variables are proaction and retroaction. Proaction refers to the material or tasks learned or performed prior to the learning of the task whose retention is measured; retroaction is the material or tasks learned or performed between the learning of a given task and the test of its retention.

Experiments in the area of verbal learning and retention, such as those by Whitely and Blankenship (11), Twining (7), Melton and Von Lackum (4), McGeoch and Underwood (3), Thune and Underwood (6), and Underwood (8), have shown that under a variety of experimental conditions both proaction and retroaction lower performance on the retention test. The three experiments reported by Underwood (8) constitute a particularly clear demonstration that performance on the retention test is inversely related to the amount of proaction and amount of retroaction. He found that the greater the number of lists of words learned prior to the learning of a given list, the less the recall score for that list, and the greater the number of lists learned between the learning and retention test of a given list or the greater the number of practice trials on one such interpolated list, the less the recall of that list.

From Underwood's and other related experiments it would seem to follow that if the retention of a task is subject to *both* proaction and retroaction at the same time, then the two effects should summate in some fashion. In fact, the retention of any task, whether learned in the laboratory or outside, is subject to interference or facilitation by both proaction and retroaction. In typical experiments conducted to date, however, one of these classes of variables has usually been held constant while the effects of changes in the other are studied. Thus, if an investigator is interested in the effects of different amounts of retroaction on retention of some material, he may use different groups of subjects,

[1]This study was supported by Grant no. APA–39 from the National Research Council to E. Tulving.

each learning a different amount of interpolated material, but all being presumably equal on variables classified as proaction. For the so-called unpractised or naïve subjects the amount of experimentally manipulated proaction is zero in such an experiment, for practised subjects the amount of experimentally manipulated proaction may be defined in terms of the practice tasks that they all are required to do prior to the experiment proper or in terms of prior similar experimental experience. The essential point is that the class of variables called proaction in a typical retroaction experiment is controlled either statistically or experimentally, but not varied systematically, and so is the class of variables called retroaction in a typical proaction experiment.

If now, under certain conditions at least, learning a list of words prior to another lowers the recall of the second list (proactive inhibition), and if learning a list of words between the learning and recall of another also lowers the recall of the first (retroactive interference), then what would happen if the retention of a given list of words is subject to *both* proaction and retroaction? Consider, for instance, the following paradigm:

Group I	Learn A	Learn B		Recall B
Group II		Learn B	Learn C	Recall B
Group III		Learn B		Recall B
Group IV	Learn A	Learn B	Learn C	Recall B

If tasks A, B, and C consist of learning comparable lists of words we could predict that the recall of the critical list B is less for Group I than for Group III (proactive inhibition) and that it is less for Group II than for Group III (retroactive inhibition). Following Melton and von Lackum (4) we could also predict that for a short retention period recall of List B is less for Group II than for Group I or that proactive inhibition is smaller than retroactive inhibition.

The question that to the best of our knowledge has not been asked by any previous experimenters and the one that was of primary interest to us in this experiment is this: How does recall of B by Group IV compare with that by Groups I and II? Are the two effects, proactive and retroactive inhibition, additive, and if so, how are the effects combined? In other words, although it produces a real tongue-twister, what is the interaction between proaction and retroaction in short-term retention? The question can, of course, be asked about different lengths of retention periods. In fact, Underwood (9) has shown that there may be some interaction between amount of proaction and amount of retroaction on the one hand and the length of retention period on the other hand in determining recall performance, and hence it is quite possible that the kind of interaction found between proaction and retroaction itself changes

with retention interval. In the present experiment, however, we limited our original question to a very short-term retention only.

We used a rather novel design to get at this problem. It is not completely free of defects, but it permits at least a rough answer to the question we asked. The design is a straightforward extension of the paradigm presented above, combined with a relatively seldom used method of learning verbal material and the method of free recall.

Method

Apparatus and Materials

Five lists of 16 words each were constructed. All words were 5- to 7-letter dissyllabic English nouns. For reasons not pertinent to this experiment the words in each list were selected from Thorndike's and Lorge's dictionary (5) according to their frequency of usage in printed English as follows: four words from among the thousand most frequently used words, four words from among the fifth thousand, four words from among the tenth thousand, and four words from the thirtieth thousand. A further restriction was that no two words in the same list should start with the same letter. The lists were varied in content on the basis of a pilot study so that all five lists were of approximately equal difficulty. The final lists are shown in Table I.

TABLE I

Lists of Words Used in the Experiment

List I	List II	List III	List IV	List V
action	buyer	accent	canard	amice
bandage	crumpet	barrack	daddy	bridle
crater	dipper	drumlin	flower	cherub
dinner	equine	finding	gambler	doctor
express	farrow	garden	hermit	ether
fiber	granite	hoyden	island	famine
gable	kitty	issue	letter	hamlet
hormone	linden	jungle	miser	lactose
legion	midden	lagoon	noodle	nature
octroi	novice	maxim	person	object
pustule	ocean	office	rumour	pilgrim
quinsy	picture	pomade	surtout	reason
satin	quarter	quillet	trochee	stamen
travel	result	treason	voter	trollop
viceroy	talker	valley	wafer	vulture
wallow	virgin	walker	zenith	waiver

For each list 16 different word orders were constructed in a systematic fashion such that each word appeared in each serial position just once and followed each other word in the list just once.

Two Gerbrands memory drums were used to present words at the rate of one word per 1½ sec. The drums and the *E* were separated from the *S* by a cardboard shield, with 2 by ½ in. slots cut to allow S to see the words appearing on the drums.

Subjects

Ss were 60 University of Toronto first-year undergraduates enrolled in the Introductory Psychology course. Their age range was 18 to 30 years, with a median age of 20 years. They were randomly assigned to 5 experimental groups of 12 Ss each.

Procedure and Design

Each *S* learned 1, 2, 3, 4, or 5 lists consecutively, depending upon the experimental group to which he was assigned, and after a 10 min. interpolated activity attempted to recall all words from all previously learned lists. The detailed procedure was as follows: The first list of 16 words was presented to *S* 8 times in succession, in a different order on each trial, and with 5 sec. between trials. *S* then attempted to write down as many words as he remembered from the list. He was given 90 sec. to do this. No restrictions were placed on the order of words in recall. *S*'s performance in this test for immediate recall, measured in terms of the number of words correctly recalled, was regarded as his learning (L) score for that list. The second list (if given) was then presented in the same way and the L-score obtained, then the third list, and so on until the number of lists prescribed for the *S*'s group had been learned. *S* was then given 10 min. to complete the Shipley-Hartford Abstraction Test, and finally was instructed to write down as many words as possible from all of the previously learned lists. The time allowed for this final recall varied according to the experimental group; for each previously learned list 90 sec. were given to *S* for recall. Again no restrictions were imposed on the order of words in recall nor were *Ss* themselves asked to identify recalled words by the ordered number of the list from which each word came. All the words recalled by each *S* were inspected by *E* and assigned to various ordered lists for that *S*. The number of words correctly recalled by *S* from each previously learned list constituted his retention (R) score for that list. In a few cases words were recorded by *S* in the final recall which were not among those previously learned by *S*. These were disregarded in scoring.

In the original learning 6 of the 12 *Ss* in each experimental group were presented with 8 of the 16 different word orders of each list, and the other half of each group were shown the remaining 8 word orders of each list. The lists were assigned to individual *Ss* in a semi-systematic fashion so that different *Ss* learned different lists and different orders of lists.

The design is illustrated in Table II. From the table it can be seen, to consider Group D as an example, that the R-score for ordered list No. 3 is subject to 2 lists as proaction (ordered lists 1 and 2) and 1 list as retroaction (ordered list 4). Similarly, Group E's recall of the second list is subject to the proactive influence of 1 list and the retroactive influence of 3 lists, and so on. The final recall of the sole list in Group A, of course, is free from experimentally controlled proaction and retroaction.

TABLE II

EXPERIMENTAL DESIGN

Group	Original learning (ordered lists)					10 min. retention interval	Final recall
A					1	Standard activity (Shipley-Hartford)	Free recall of words in all previously learned lists
B				1	2		
C			1	2	3		
D		1	2	3	4		
E	1	2	3	4	5		

The independent variables, proaction and retroaction, are thus measured in terms of the number of comparable lists learned prior to and following any given list. This design makes it possible to obtain measures of retention that are subject to various combinations of amounts of proaction and retroaction. Table III shows the categories of treatment combinations into which fall the data from final recall, and indicates which of the 5 groups contribute the data for each combination.

TABLE III

GROUPS CONTRIBUTING DATA FOR EACH COMBINATION OF AMOUNTS OF RETROACTION AND PROACTION

Amount of proaction (number of lists)	Amount of retroaction (number of lists)				
	0	1	2	3	4
0	A_1*	B_1	C_1	D_1	E_1
1	B_2	C_2	D_2	E_2	
2	C_3	D_3	E_3		
3	D_4	E_4			
4	E_5				

*Capital letters denote experimental groups, subscripts refer to the number of ordered lists.

RESULTS

The data included the number of words recalled immediately after the eight learning trials for a list (Learning or L), and the number of words from each list in final recall (Retention or R). These data are summarized in Table IV and Figures 1 and 2.

Table IV shows means and standard deviations of L-scores for each ordered list for each group. It is clear from the column means that there are no systematic differences between L-scores for different ordered lists.

TABLE IV

NUMBER OF WORDS RECALLED IMMEDIATELY AFTER EIGHT LEARNING TRIALS (L-SCORES): MEANS AND STANDARD DEVIATIONS

Group		Ordered lists				
		1	2	3	4	5
A	M	12.9				
	S.D.	1.8				
B	M	11.5	11.7			
	S.D.	2.3	3.0			
C	M	11.6	11.3	12.8		
	S.D.	1.5	0.96	1.9		
D	M	10.4	11.3	10.9	11.7	
	S.D.	2.2	2.6	2.8	2.1	
E	M	11.3	11.7	11.6	11.4	11.7
	S.D.	1.3	2.1	1.9	2.8	1.9
Column means		11.5	11.5	11.8	11.6	11.7

That is, there are no resultant effects due to warm-up, practice, or transfer. It is also evident that there are no substantial differences between L-scores for different groups.

The retention data were of primary interest in the experiment. Figure 1 shows mean R-scores as a function of retroaction for various amounts of proaction, and Figure 2 illustrates the same data with proaction on the abscissa and retroaction as the parameter. The analysis of these data was complicated by the facts that (1) the effects of the two independent

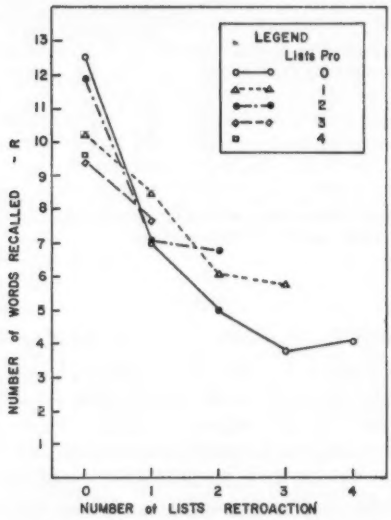

FIGURE 1. Ten-minute recall as a function of amount of retroaction. Each point represents the mean number of words recalled by 12 Ss under different amounts of retroaction (abscissa) and different amounts of proaction (parameter).

variables and their interaction are partially confounded with differences between subjects, and (2) the 5 × 5 proaction-retroaction table is incomplete. These difficulties are evident upon examination of Table III.

A preliminary analysis was performed upon the 3 × 3 table composed of the data from the combinations of zero, one, and two lists proaction and retroaction, assuming a completely orthogonal design. Under this assumption, the statistical model is not quite correct for the design used, but the tests are conservative. An analysis of covariance was used on

R with L as a covariant. The test for parallelism of individual regression lines indicated that it was reasonable to assume no interaction between L and the experimental treatments ($F = 1.72$, 8 and 90 df, $p > 0.10$). The covariance analysis showed that the interaction between proaction and retroaction was not significant ($F = 1.81$, 4 and 98 df, $p > 0.10$). Proaction was also not significant ($F = 0.73$, 2 and 98 df), but the hypothesis of no retroactive effect could not be rejected ($F = 36.4$, 2 and 98 df).

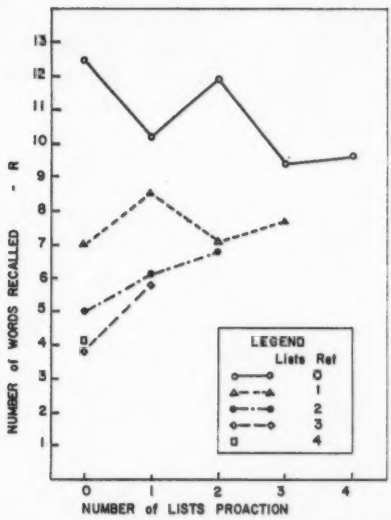

FIGURE 2. Ten-minute recall as a function of amount of proaction. Each point represents the mean number of words recalled by 12 Ss under different amounts of proaction (abscissa) and different amounts of retroaction (parameter).

The graph of Figure 2 suggests that the inclusion of the data missing from the preliminary analysis might result in the rejection of the null hypothesis of no interaction between proaction and retroaction. That is, for the sample used in this experiment, R decreases with increasing proaction for zero retroaction, but for one and two lists retroaction, proaction actually facilitates recall. However, individual analyses of covariance of R with L as covariant for zero, one, two, and three lists retroaction resulted in acceptance of the null hypothesis in each case

($F = 2.14$, 4 and 54, df, $p > 0.05$; $F = 0.932$, and 43 df; $F = 1.63$, 2 and 32 df, $p > 0.10$; $F = 1.30$, 1 and 21 df, $p > 0.25$; respectively). This means that there is not sufficient evidence to reject the hypothesis that all curves in Figure 2 are parallel to one another and to the abscissa.

Thus it can be concluded that increasing retroaction produces increasing decrement in 10 min. recall, whereas increasing proaction has no demonstrable effect.

Discussion

If we interpreted our results strictly in keeping with the practices of statistical zealots we would be forced to the conclusion that under the conditions of this experiment retroaction has a pronounced effect on 10 min. retention and that proaction has no effect. This would mean then that what affects retention is what has been learned after the learning of the list whose retention is measured, but not what has been learned prior to the learning of the designated list. We doubt whether anyone would have much quarrel with the statement that retroactive inhibition is a phenomenon which can be demonstrated under a wide range of experimental conditions, including those of the experiment reported here. For three reasons, however, we are somewhat reluctant to make a strong statement about proactive effects in this experiment. First, our statistical model, for reasons that we pointed out above, permits only conservative tests of the null hypothesis. In any case, when accepting the null hypothesis we might have made a Type II error. Second, the top curve in Figure 2, showing R as a function of proaction with retroaction absent, strongly suggests proactive inhibition. The computed F for this curve ($F = 2.14$, 4 and 54 df) is significant at the "Oh Shucks!"[2] level, and the interaction F in the preliminary analysis mentioned above also approached significance at the same level, suggesting interaction between proaction and retroaction. Third, other experimenters have clearly demonstrated proactive inhibition under somewhat different experimental conditions (**1, 4, 8**) and we like to believe in empirical evidence Indeed, Underwood (**10**) has recently suggested that proaction plays a much more important role in forgetting than is commonly believed. It is possible, of course, that the absence of any proactive effects in the present experiment is related to the differences between the procedures used by previous experimenters and the design used in this experiment. If this is the case it becomes important to isolate and specify the variable or variables responsible for such differences. In view of the state of theories in this field at present such a problem would seem to be primarily an empirical one.

[2]Term used by G. A. Kelly, *Ann. Rev. Psychol.*, 1958, 9, p. 325.

From a look at the data in Figure 2, it is tempting to suggest that there is interaction between proaction and retroaction and that our failure to obtain a significant F for such interaction is related to large variability between the subjects. If we discard, at least for the present, the somewhat improbable possibility that proaction actually facilitates recall when combined with relatively greater amounts of retroaction, although it seems to be suggested by the sample curves for two and three lists of retroaction in Figure 2, then it would seem that proactive inhibition decreases with increasing amounts of retroaction. In such a case it would also follow logically that relative retroactive inhibition should be greatest with minimum amounts of proaction and that it decreases with increasing amounts of proaction. An evaluation of these hypotheses has to wait for further experimental work, since the evidence for them in this study was insufficient. If these hypotheses are supported by empirical evidence it would provide the theorists in this area with another whetstone on which to sharpen their tools.

There is relatively little that we would like to say about the effects of retroaction on retention found in this experiment. Our findings simply confirm the results from many previous studies that have demonstrated retroactive inhibition. The curve for zero amount of proaction in Figure 1 seems to be flattening off and approaching an asymptote, corroborating findings by Underwood (**8**) and Briggs (**2**).

The design that we have used in this experiment is not free from defects, but it has permitted us to look at the combined effects of proaction and retroaction and it can be considered a useful tool for extending our understanding of the phenomenon of retention and forgetting. An essential element in this design is the learning task given to the subject where he is required to learn to make individual responses without many restrictions upon stimuli to which these responses are appropriate. This makes it possible to ask for "free" recall of previously learned responses in the subsequent retention test and permits responses to be emitted in the order of their relative strengths. Although it is possible that the subject's earlier responses in the retention test may interfere with other, yet-to-be-made responses and thus introduce another source of error, the difficulty is not as serious with the method used here as it would be with the more traditional methods of serial anticipation and paired associates which have been used almost exclusively by previous investigators in studying proactive and retroactive effects on retention. The order in which responses are made by the subject in free recall is probably influenced by prior experimental conditions and by the independent variables used in the experiment. The order of recalled words in the first relearning trial with the method of

serial anticipation or paired associates is, however, also influenced by the order in which the experimenter presents appropriate stimuli to the subject, which may constitute another source of variability. With these more orthodox methods another difficulty arises in connection with intra-list and inter-list intrusions. How should these be treated? Is an intrusion comparable to no response, indicating subliminal strength of that response? In what sense is an intrusion retained or forgotten? The method used in this experiment sidesteps these problems.

The main difficulty inherent in the design used has already been pointed out above: statistical analysis is difficult because of lack of an appropriate model. It is possible, of course, to eliminate partial confounding between subjects and the effects of the two independent variables by using data for one ordered list only from each group of subjects. Thus there should be two groups of subjects instead of a single Group B, three groups instead of Group C, and so on, in the present type of experiment. Eventually this may have to be done. At present, however, there are many more important sources of error to worry about and hence such a "pure" design may seem wasteful of experimenters' and subjects' time. In this experiment we should have used three times as many subjects as we did to get the same amount of data and a "pure" design.

Summary

In this experiment Ss learned lists of disparate English nouns by a modified method of serial presentation. The learning of each list was preceded and followed by the learning of other comparable lists. After a 10 min. period of interpolated activity retention scores for previously learned lists were obtained. The results showed that the number of lists learned prior to the learning of any particular list (amount of proaction) did not affect retention. Recall scores, however, were found to be inversely related to the number of lists that had been learned after a particular list (amount of retroaction), thus demonstrating retroactive inhibition. There was not sufficient evidence for accepting the hypothesis of interaction effects between proaction and retroaction although such interaction was suggested by the data.

REFERENCES

1. Atwater, S. K. Proactive inhibition and associative facilitation as affected by degree of prior learning. *J. exp. Psychol.*, 1953, **46**, 400–404.
2. Briggs, G. E. Retroactive inhibition as a function of the degree of original and interpolated learning. *J. exp. Psychol.*, 1957, **53**, 60–67.
3. McGeoch, J. A., & Underwood, B. J. Tests of the two-factor theory of retroactive inhibition. *J. exp. Psychol.*, 1943, **32**, 1–16.
4. Melton, A. W., & Von Lackum, W. J. Retroactive and proactive inhibition in retention: Evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **54**, 157–173.

5. THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words.* New York: Teachers College, 1944.
6. THUNE, L. E., & UNDERWOOD, B. J. Retroactive inhibition as a function of degree of interpolated learning. *J. exp. Psychol.*, 1943, **32**, 185–200.
7. TWINING, P. E. The relative importance of intervening activity and lapse of time in the production of forgetting. *J. exp. Psychol.*, 1940, **26**, 483–501.
8. UNDERWOOD, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychol. Monogr.*, 1945, **59**, no. 273.
9. UNDERWOOD, B. J. Retroactive and proactive inhibition after 5 and 48 hours. *J. exp. Psychol.*, 1948, **38**, 29–38.
10. UNDERWOOD, B. J. Interference and forgetting. *Psychol. Rev.*, 1957, **64**, 49–60.
11. WHITELY, P. L., & BLANKENSHIP, A. B. The influence of certain conditions prior to learning upon subsequent recall. *J. exp. Psychol.*, 1936, **19**, 496–504.

MOTIVATIONAL VARIABLES IN SECOND-LANGUAGE ACQUISITION[1]

ROBERT C. GARDNER AND WALLACE E. LAMBERT

McGill University

MOST RESEARCH on second-language acquisition has been concerned with the measurement of an "ability for languages," the assumption being that achievement is largely due to a linguistic aptitude. However, when measures of aptitude are correlated with grades in language courses, the validity coefficients show considerable variability from situation to situation even with tests developed through factor analytic methods (**2**), suggesting that variables other than linguistic aptitude are involved. Researchers have mentioned that motivation and interest probably play important roles in second-language acquisition (**4, 5, 8, 16, 18**), but perhaps because of difficulties in measuring them, these aspects have not been given systematic attention.

Theoretical attempts to explain how the child learns his first language have emphasized a particular type of motivation. Mowrer's theory (**11**) suggests that language acquisition is motivated by a desire to be like valued members of the family and, later, of the whole linguistic community. Ervin (**6**) has extended this view, suggesting that emotional dependence or respect for another individual may account for some instances of marked success in second-language achievement. Support for this extended interpretation has been found in recent studies where (*a*) fluency in the second language was found to depend upon an active interest in members of the other linguistic community (**19**); (*b*) both extent of bilingualism (**12**) and advanced level of language study (**7**) were associated with lower *F*-scores; (*c*) dominance in the acquired language was related to personal dissatisfactions with one's own group (**9**). In line with this latter finding, adults electing to study French conversational courses scored higher on Srole's Anomie Scale (**14**) than students enrolled in the more grammatical courses, suggesting that adults

[1]This research was supported by the Canadian Defence Research Board, Grant no. D77–94–01–10 and by a Ford Foundation Grant to D. O. Hebb. We are grateful to both Professor G. A. Ferguson (McGill) and Professor J. B. Carroll (Harvard) for their advice on statistical analysis. In particular, we want to thank J. T. McIlhone of the Catholic School Board of Montreal for his co-operation and kindness in permitting us to carry out this study.

dissatisfied with their position in their own cultural group are seeking to learn, as rapidly as possible, those aspects of the other language which will allow them to become members of a new group (7).

It is our contention then that achievement in a second language is dependent upon essentially the same type of motivation that is apparently necessary for the child to learn his first language. We argue that an individual acquiring a second language adopts certain behaviour patterns which are characteristic of another cultural group and that his attitudes towards that group will at least partly determine his success in learning the new language. Our use of attitude as a motivational construct presupposes an intention on the part of students to learn the language with various aims in mind, and to pursue these aims with varying degrees of drive strength. Our test battery consequently included indices of motivational intensity and orientation. The "Orientation Index" classifies purposes in one of two ways: "integrative," where the aim in language study is to learn more about the language group, or to meet more and *different* people; "instrumental," where the reasons reflect the more utilitarian value of linguistic achievement. The "Motivational Intensity Scale" measures the amount of effort and enthusiasm students show in their attempt to acquire the language.

The present research was designed to determine the comparative importance of linguistic aptitude and certain motivational variables[2] in learning a second language. The intercorrelations of a battery of tests administered to high school students were analysed to determine the factorial structure and to see how the criterion variable, skill in speaking and understanding French, related to this structure.

Method

Subjects

Forty-three male and thirty-two female English-speaking Grade XI high school students acted as Ss. They had already completed an average of about 7 years of formal training in French.

Materials

Scores on the following tests were obtained:

(1) Achievement ratings: The French instructor of each class was asked to rate the students on each of two attributes: oral skills (ability to imitate French word sounds) and aural comprehension (understanding of spoken French). Ratings were made on a 5-point scale from poor (1) to excellent (5). Since these 2 ratings were highly correlated, they were combined, then standardized, to arrive at the achievement rating, Variable 1.

Variables 2–6 are the five sub-scales from Carroll's "Psi-Lambda Foreign Language

[2]Copies of these scales can be obtained from the authors at McGill University.

Aptitude Battery" (3). They include: (2) number learning, (3) phonetic script, (4) spelling clues, (5) words in sentences, and (6) paired associates.

Variables 7 and 8 are two subscales from the A.C.E. College Entrance Examination, 1943 edition. Those used were: (7) same-opposites, and (8) verbal analogies.

(9) Orientation index: Ss were presented with four alternative reasons for studying French and asked to rank them as to their personal relevance. They indicated that a knowledge of French would: (1) be useful in obtaining a job, (2) be helpful in understanding the French-Canadian people[3] and their way of life, (3) permit meeting and conversing with more and varied people, and (4) make one a better-educated person. A fifth alternative, "any other personal reason," was supplied in case an individual thought the alternatives given were inadequate. Ss who ranked either alternative (2) or (3) as most relevant were classified as "integratively oriented." Those choosing alternatives (1) or (4) were classified as "instrumentally oriented." Those choosing alternative (5) were not classified.

(10) Attitude scale: Twenty positively worded statements about French-Canadians were presented and S was asked to indicate his degree of disagreement or agreement on a 7-point scale. A sample item is: "The French-Canadian has every reason to be proud of his race and culture."

(11) Motivational-intensity scale: Eight multiple choice statements were designed to measure intensity of motivation in terms of: work done for assignments, preference for and comparative ease of the course, opportunities taken to improve speaking and reading French, future intentions to study or make use of the language, and importance attributed to a knowledge of the language. Ss were asked to choose, in each case, the alternative most descriptive of themselves.

(12) California *F*-scale: Twenty-five items from forms 45 and 40 (1).

(13) Sex: This was included as a separate variable to account for sex differences.

(14) Audience sensitivity scale: This scale (13) is a measure of the typical anxiety responses made by the individual in a variety of social situations. It was included in the present battery because it was thought that such anxiety might deter a student from practising the second language.

Results and Discussion

Product-moment intercorrelations were computed for twelve of the tests, and the correlations of these twelve with variables 9 and 13 were computed as biserial correlation coefficients. The correlation between variables 9 and 13 is a *phi*-coefficient. The correlation matrix is given in Table I.

[3]The term "French-Canadian" instead of "French" was used to make Ss think in terms of the language group with which they come into contact in Montreal. Although the courses use Standard French with little or no reference to any French-Canadian modifications, the French learned in school can be used with no difficulty by English-speaking students in the province of Quebec. It may be true that a few Ss were motivated to learn French because of its possible value to them in the European French community and these Ss will tend to reduce the relationships of the orientation index and attitude scale (see below) with other measures.

TABLE I

CORRELATION MATRIX*

	1	2	3	4	5	6	7	8	9†	10	11	12	13‡	14
1. Achievement ratings	—	.22	.28	.28	−.01	.18	.16	.42	.34	.10	.40	.08	−.06	−.08
2. Number learning		—	.33	.04	.10	.25	.22	.41	−.24	−.18	.13	.03	−.18	−.07
3. Phonetic script			—	.47	.31	.58	.34	.44	.21	.13	.22	.15	.28	.05
4. Spelling clues				—	.29	.43	.48	.40	−.04	.23	.20	−.02	.08	−.08
5. Words in sentences					—	.13	.30	.21	.17	.15	.04	−.07	.22	.03
6. Paired associates						—	.37	.53	.08	.32	.22	.13	.43	−.37
7. Same-opposites							—	.56	−.14	.16	.01	−.14	.24	.08
8. Verbal analogies								—	.14	−.05	.12	−.19	.12	.00
9. Orientation index									—	.42	.44	−.02	.01	−.04
10. Attitude scale										—	.44	.27	.28	−.03
11. Motivational-intensity											—	.24	.04	−.10
12. California *F*-scale												—	.31	.04
13. Sex													—	.14
14. Audience sensitivity scale														—

*It was not possible with the testing time given us to have *Ss* complete all parts of the Foreign Language Aptitude Test (Variables 2–6). A different test of the battery was omitted for each school class. Consequently, the total *N* is reduced by approximately 20 for all correlations involving these variables.

†Biserial correlations are positive if the integratively oriented *Ss* obtained a higher mean score on the continuous variables than did those instrumentally oriented.

‡Biserial correlations are positive if the females scored higher on the tests.

Four factors (see Table II) were extracted by Thurstone's centroid method (17). Graphical representations of these factors plotted two at a time indicated that no meaningful rotation could be obtained with the fourth factor, and this factor was subsequently omitted in the rotations. Examination of the plots of the first three factors suggested that an orthogonal solution was warranted, and that little could be gained with oblique rotations. The rotated factor loadings are given in Table III.

TABLE II

CENTROID FACTOR MATRIX

Measures	I	II	III	IV	Communalities
1. Achievement ratings	.47	.11	−.46	.20	.49
2. Number learning	.26	−.39	−.27	.30	.39
3. Phonetic Script	.70	−.17	.10	.13	.55
4. Spelling clues	.56	−.26	.18	−.16	.44
5. Words in sentences	.35	−.13	.16	−.09	.17
6. Paired associates	.76	−.22	.11	−.12	.65
7. Same-opposites	.50	−.40	.20	.10	.46
8. Verbal analogies	.60	−.41	−.24	.22	.64
9. Orientation index	.30	.46	−.26	−.25	.43
10. Attitude scale	.45	.48	.24	−.30	.58
11. Motivational-intensity	.50	.46	−.30	−.08	.55
12. California *F*-scale	.17	.40	.24	.26	.31
13. Sex	.34	.11	.55	.17	.46
14. Audience sensitivity scale	−.13	.09	.19	.42	.24

TABLE III

ROTATED FACTOR MATRIX
(ORTHOGONAL)

Measures	I	II	III	IV	Communalities
1. Achievement ratings	.30	.55	−.23	.20	.49
2. Number learning	.45	.00	−.31	.30	.39
3. Phonetic script	.66	.21	.24	.13	.55
4. Spelling clues	.60	.20	−.08	−.16	.43
5. Words in sentences	.36	.02	.20	−.09	.17
6. Paired associates	.73	.20	.24	−.12	.64
7. Same-opposites	.64	−.10	.18	.10	.46
8. Verbal analogies	.73	.15	−.19	.22	.64
9. Orientation index	−.05	.60	.03	−.25	.43
10. Attitude scale	.05	.47	.52	−.30	.59
11. Motivational-intensity	.10	.73	.06	−.08	.55
12. California *F*-scale	−.12	.26	.41	.26	.31
13. Sex	.19	.00	.62	.17	.45
14. Audience sensitivity scale	−.16	−.10	.16	.42	.24

The factor analysis indicates that two independent factors are related to achievement in French. Factor I has its highest loadings on Tests 1–8. Test 1 is the criterion, suggesting that the others measure some component related to achievement in French. Tests 7 and 8 are the two indices of verbal intelligence while Tests 2–6 are the subscales from the language aptitude battery. Factor I, therefore, is clearly a *linguistic aptitude* factor.

It should be noted that the high loadings of the verbal intelligence tests on this factor lend support to the findings of Wittenborn and Larsen (**20**) that intelligence and language aptitude are factorially similar.

Factor II derives its highest loadings from Tests 1, 9, 10, and 11. Since the criterion is substantially loaded on this factor, the latter three variables must be considered important for the successful acquisition of a second language. Examination of these three variables suggests that Factor II should be defined as a *motivation* factor. It should be emphasized however that this denotes a motivation of a particular type, *characterized by a willingness to be like valued members of the language community.*

These results indicate, therefore, that two components are related to second-language achievement. That this is not an artifact of the factorial rotations is substantiated by examining those tests which produce the maximum multiple correlation with the criterion. Four tests (8, 11, 9, and 5) were selected by the Wherry-Doolittle test selection technique (**15**) which yielded the maximum multiple correlation of 0.558. These tests are equally divided between the two factors indicating the bidimensional character of the variables associated with second-language achievement.

In an exploratory study such as this, it is not advisable to pay too much attention to relatively small differences in the magnitude of specific factor loadings. The important result is that two factors, not one as previously presumed, are associated with second-language achievement. Furthermore, the variables constituting the "new" factor have an empirical as well as a theoretical basis. The significant positive correlation between the orientation index and achievement in French indicates that the integratively oriented students are generally more successful in acquiring French than those who are instrumentally oriented. Further, the students with the integrative orientation have more favourable attitudes towards members of the French group and are more strongly motivated to acquire their language. Although we favour the view that a strong motivation to learn a second language follows from a desire to be accepted as a member of the new linguistic community (cf. **12**), the correlation technique does not permit us to be completely certain of such an interpretation. However, with the demonstration presented here of the importance of the "motivational factor," further study of the integrative orientation can incorporate experimental procedures to determine the actual causal sequence of events.

Summary

Montreal high school students studying French as a second language completed a battery of tests including measures of linguistic aptitude, verbal intelligence, and various attitudinal and motivational characteristics. Analysis of the intercorrelations

of these tests yielded two orthogonal factors equally related to ratings of achievement in French: a "linguistic aptitude" and a "motivational" factor. It was also found that maximum prediction of success in second-language acquisition was obtained from tests of: verbal intelligence, intensity of motivation to learn the other language, students' purposes in studying that language, and one index of linguistic aptitude.

REFERENCES

1. Adorno, T. W., Frenkel-Brunswik, Else, Levinson, D. J., & Sanford, R. N. *The authoritarian personality.* New York: Harper, 1950.
2. Carroll, J. B. Summary of validity coefficients–foreign language aptitude battery. Harvard University, 1956 (mimeographed).
3. Carroll, J. B., & Sapon, S. M. Supplement 1 to examiner's manual–Psi-Lambda foreign language aptitude battery. Harvard University, 1956 (mimeographed).
4. Carroll, J. B. A factor analysis of two foreign language aptitude batteries. *J. gen. Psychol.*, 1958, **59**, 3–19.
5. Dunkel, H. B. *Second-language learning.* Boston: Ginn, 1948.
6. Ervin, Susan. Identification and bilingualism. Harvard University, 1954 (mimeographed).
7. Gardner, R. C. Social factors in second-language acquisition. Unpublished Master's thesis, McGill University, 1958.
8. Henmon, V. A. C. Prognosis tests in the modern foreign language. *Publications of the American and Canadian Committees on Modern Foreign Languages*, 1929, **14**, 3–31.
9. Lambert, W. E. Measurement of the linguistic dominance of bilinguals. *J. abnorm. soc. Psychol.*, 1955, **50**, 197–200.
10. Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. Evaluational reactions to spoken languages. *J. abnorm. soc. Psychol.*, (in press).
11. Mowrer, O. H. *Learning theory and personality dynamics.* New York: Ronald, 1950.
12. Newcomb, T. M. *Social psychology.* New York: Dryden, 1950.
13. Paivio, A. V., & Lambert, W. E. Measures and correlates of audience anxiety (stage fright). *J. Pers.*, 1959, **27**, 1–17.
14. Srole, L. Social dysfunction, personality and social distance attitudes. (Paper read before American Sociological Society, 1951 National Meeting, Chicago, Illinois.)
15. Stead, W. H., Shartle, C. L., *et al. Occupational counseling techniques.* American Book Company, 1940.
16. Symonds, P. M. A modern foreign language prognosis test. *Publications of the American and Canadian Committees on Modern Foreign Languages*, 1929, **14**, 91–126.
17. Thurstone, L. L. *Multiple factor analysis.* Chicago: Univer. Chicago Press, 1947.
18. Todd, J. W. The psychological fundamentals of linguistic achievement. *Publications of the American and Canadian Committees on Modern Foreign Languages*, 1929, **14**, 129–172.
19. Whyte, W. F., & Holmberg, A. R. Human problems of U.S. enterprise in Latin America. *Human Organization*, 1956, **15**, No. 3, 1–40.
20. Wittenborn, J. R., & Larsen, R. P. A factorial study of achievement in college German. *J. educ. Psychol.*, 1944, **35**, 39–48.

AN OBJECTIVE EVALUATION OF BRIEF GROUP PSYCHOTHERAPY ON DELINQUENT BOYS[1]

B. ROGER PHILIP AND HELEN E. PEIXOTTO

Assumption University of Windsor *Catholic University of America*

THERE IS AMPLE EVIDENCE from the clinical standpoint that group psychotherapy, even of a relatively brief nature, produces beneficial personality changes. Often, however, the evidence for improvement is too dependent upon the clinical judgment of a few observers to be fully convincing to experimentally minded critics.

Criteria discussed by Zubin (7), Frank (2), and Burchard *et al.* (1) have been utilized by several recent investigators who have attempted to evaluate group psychotherapy. There is a growing tendency to use objective or semi-objective methods in such studies as is evidenced by the work of Jones and Peters (4) and Muench (5). More pertinent to this paper are the studies of Teuber and Powers (6) and especially Gerston (3) who applied these methods to the study of delinquent boys.

Teuber and Powers studied 650 boys, aged 6–10 years, who were likely to become delinquent. The subjects were split into two groups, treatment and control, whose members were individually matched on a number of variables. The treatment varied from friendly "big brother" relationships to formal psychiatric interviews; all treatment was withheld from the control group. The effects of the therapy were evaluated by the number of offences committed by both groups from the beginning of the treatment until three years later. Using this criterion it was found that 96 boys of the treatment group committed 264 offences, and 92 boys of the control group committed 218 offences.

Gerston's subjects were 44 boys, aged 13 to 16 years, who were divided into matched experimental and control groups, both of whom were given a battery of tests at the beginning and at the end of the experiment. The experimental group participated in 22 therapy sessions of one hour each, once a week. Slawson's method of group therapy was used for the first quarter of the treatment, and handicraft therapy thereafter. When the groups were compared at the end of the therapy sessions there was a gain on the Stanford Achievement test of 18 months for the experimental group and of three months for the control group; the difference was

[1]This study was supported by a grant from the Canada Council. Appreciation is due to the officials of St. John's Training School, Uxbridge, for their co-operation.

significant at the .01 level. The experimental group also showed greater gains compared with the control group on emotional and social maturity as measured by the Haggarty-Olsen-Wickman scale, and more R, M, FC, CF, H and P on the Rorschach.

In the study reported here an attempt was made to measure objectively certain personality changes resulting from brief group psychotherapy with delinquent boys. Boys of this type are suspicious of tests based upon questionnaires and sometimes they purposely distort their replies. Accordingly, it was thought that some form of projective technique might be more satisfactory. There is available a modified form of the TAT, devised by the Commission on Human Resources and Advanced Training, which is objectively scored and which does not patently disturb the subject. This test, in a slightly modified form, was used in this investigation and will be described later.

Method

Subjects

The Ss who took part in this investigation were 86 delinquent boys, 14 to 16 years of age, all of whom had been court committed, for minor misdemeanours, to St. John's Training School, Uxbridge, Ontario.

They were divided into 2 groups, an experimental group and a control group, of 43 boys each who were matched for age, grade, intelligence, and Hostility scores on the TAT. The mean values for these variables for all Ss are shown in Table I, which

TABLE I

Comparison of the Experimental and the Control Groups

Group	*N*	Age in months	I.Q.	Grade	Months in school	No. of repeaters	Hostility score
Experimental	43	174	99	7.9	4.7	4	63.2
Control	43	176	96	7.5	5.1	4	62.8

also gives the average length of time the Ss of both groups had been in the School, and the number of repeaters in each group, that is, those who were in the School for a second time. All Ss had passed a Gates Reading Survey at the Grade VII level or better. The experimental group was subdivided for therapy purposes into 4 sections. Two of the sections (N=19) had been selected in an individual initial interview given by the senior author, where they had shown evidence of marked hostility, often by the angry and aggressive statements they made against their parents, or their siblings, their companions, the police, and the school authorities. Some had not satisfactorily adjusted to the school programme. The mean score for these two sections on the TAT Hostility scale was 65.5. A control group, many of

whose members also were hostile, was chosen to match this hostile experimental group, and the mean score of this group on the Hostility scale was 65.6. The other two sections of the therapy group (N=24) had been chosen by one of the officials of the school on the basis of the age and grade of the boys. The mean score of the one group on the Hostility scale was 61.3, and that of the matched control 60.5.

Therapy

The 4 experimental sections each took part in 10 group therapy sessions of 1 hr. each, conducted by the senior author. These sessions were spread over 2½ weeks. At the beginning of each session a problem dealing with interpersonal relations was offered for discussion. Frequently the problems were selected by the participants in the discussion, but they were mostly the same for all groups. All Ss were encouraged to talk freely and were assured that no restrictions would be placed on what they wanted to say. When free discussion lagged, questions bearing on the problem were directed to those who had not actively participated in the discussion. Since about half of the therapy participants had been selected chiefly because they had manifested a good deal of hostility, an effort was made by the moderator to reduce this hostility by selecting for discussion topics which dealt with the problems of the interpersonal relations that seemed to disturb the Ss, such as their feelings towards the police and teachers. No attempt was made by the moderator to repress the angry and often unjust statements which were made. Later in the session some of the discussants themselves corrected exaggerated statements or offered reasonable explanations for the conduct of others. Rather infrequently the moderator asked questions or made statements, the purpose of which was to clarify the feelings of the Ss.

To mask the real purpose of the experiment and to encourage all Ss to talk freely and clearly, they were told that the sessions were being held for speech training purposes, but that at the same time they would have an opportunity to discuss their problems. The emphasis upon speech training did induce the subjects to talk clearly, and in most cases, freely. All, however, were aware, at least in a vague way, of the therapeutic nature of the sessions and some few individuals maintained a suspicious attitude throughout the sessions.

To emphasize the speech training purpose, a tape recorder and microphone were placed in full view of the subjects. They were assured that no tape would be played outside the sessions unless permission had been given to do so by all the participants. Furthermore, they were told not to discuss outside the therapy room anything that took place during the sessions. This precaution was taken to avoid disparaging comments by other boys of the Training School. The Ss were assured that their statements would be held in full confidence—a promise which was faithfully kept.

At the conclusion of every session the tape was played back to the group for a two-fold purpose. First, it enabled the Ss to hear their own voices and to note the clarity of their enunciation, thus helping the moderator to make more effective comments on their speech and inducing the Ss to participate more willingly and actively in the discussion. Secondly, the replaying of the tape served as a feedback, causing Ss to note far more critically and attentively what they had said in the discussions. Frequently comments were heard during the playback, such as "Did I say that? It wasn't quite what I meant." Often Ss revoked the rash generalizations they had made during the session and became willing to accept a more lenient attitude towards persons to whom they had previously reacted rather heatedly. Although the playing of the tape took considerable extra time it was felt that it increased the therapeutic effect of the sessions.

The Revised TAT

To evaluate some factors in the personality of Ss a form of the TAT, devised by the Commission on Human Resources and Advanced Training,[2] was used in a slightly modified fashion. This revision of the test uses 24 pictures of the TAT for each of which there are 5 descriptive sentences, designed to elicit feelings of Hostility, Need for Dependency, Blandness, Insecurity, or Guilt. Typical of the statements are those for picture 7BM, together with the feelings they are supposed to elicit:

(*a*) The boy fears that he will not be able to live up to what his father expects. (Insecurity)

(*b*) The boy is sorry because he let his father down. (Guilt)

(*c*) The boy is thinking of what his father did when he was a boy. (Blandness)

(*d*) The boy is determined to be as good a man as his father. (Need for dependency)

(*e*) The boy is mad at his father for telling him that he will be punished. (Hostility)

S is asked to mark 1 in front of the sentence he thinks best fits the picture, to mark 2 in front of the sentence which fits next best, and so on for each of the 5 sentences. All sentences are marked from 1 to 5 and Ss are allowed 1 min. for each picture. The scores are weighted by giving 8 for each first choice, 6 for the second choice, 4 for the third, 2 for the fourth and 0 for the fifth choice. The weighted scores for each of the 5 variables is totalled for all 24 pictures and is the measure of Hostility, Guilt, etc., evidenced by the Ss.

Definitions of these variables are given by the Committee on Human Resources and Advanced Training. It is conceded that the labelling of the responses as Hostility, Guilt, etc., does not indicate that these responses correspond to the true sentiments of the Ss. A sentiment like Hostility can vary enormously in intensity and complexity, and can be directed to a vast variety of objects, with consequent widely divergent personality implications. Hostility in the delinquent may be manifested for a fleeting moment or may be of long-standing duration. Certainly, any accurate appraisal of the intensity and gravity of such personality traits would require a searching and sensitive clinical appraisal. However, the tendency to oversimplify complicated concepts in the interests of clarity is common in psychological research and is not confined to this modified TAT. Nevertheless, the limitations of our personality terms must be borne in mind. What we here call Hostility, Guilt, etc. is restricted to the types of response to the various sentences describing each picture. We merely attempt to show that for the population here investigated the frequencies of the responses so labelled may be changed in certain predictable ways as a result of 10 therapy sessions, and that these changes may be objectively measured.

Since the wording of the original sentences was somewhat too difficult for our Ss, the sentences were reworded, keeping as closely as possible to the original meaning. Care was taken that all Ss fully understood the meaning of the sentences. In reply to questions from the Ss as to the purpose of the test, they were told that they would be given a test of reading comprehension based upon a description of pictures. They seemed to accept this explanation and many of the Ss asked the experimenter if they had passed the reading test. The TAT was given to groups of 24 Ss using standard pictures which were rotated among the group in such a way that, while the order of the presentation was constant, every S started with a different picture.

The tests were scored on the 5 personality traits. Two points must be emphasized:

---

[2]Privately circulated for research purposes by the Commission on Human Resources and Advanced Training, Washington, D.C. This Commission ceased to function in 1954.

first, the marking scheme is such that there is a fixed range of scores, so that when the test is repeated an increment in the score on some traits necessarily entails a decrement in the scores of the remaining traits. Secondly, the average score, as given in the published norms, varies from trait to trait. In Table II the average scores made by the 86 Ss of this experiment are compared with norms for 992 high school boys

TABLE II

COMPARISON OF SCORES ON INITIAL TAT

Group	*N*	Hostility	Achievement	Blandness	Insecurity	Guilt
Norms	992	66.3	115.1	99.8	100.3	101.0
Experimental	43	63.2	119.3	95.9	102.7	99.8
Control	43	62.8	115.2	98.2	103.8	100.0

published by the Commission on Human Resources and Advanced Training. It may be seen that the average score for Hostility is quite low as compared with that for Achievement, that the averages for our Ss do not greatly differ from the published norms, and that the experimental and control groups are fairly well matched.

Experimental Design

As has been pointed out, the experimental group consisted of Ss, many of whom had been selected for therapy on the basis of their hostility. The control group, many of whom were also hostile, resembled the experimental group in TAT score, age, intelligence and grade, and were, like them, undergoing training at the school. They differed from the experimental group chiefly in the fact that they had participated in no group therapy. The need for a control group was based upon the fact that the training given at the school is a form of therapy, "milieu therapy," and must be taken into account in any evaluation of personality changes.

All Ss were given the modified TAT about a week before the therapy began and immediately after the last therapeutic session, a lapse of some 3 or 4 weeks.

Since most of the therapy was directed at diminishing hostility in the Ss, it was hypothesized that when the first and second TAT's were compared, the experimental group would show a reduction in Hostility. It was also postulated that there would be a smaller reduction in Insecurity and Guilt. Since the scoring scheme of this test is such that a diminution of the scores for some traits is accompanied by a rise in the scores of others, it was further postulated that there would be an increase in the Dependency and Blandness scores. Owing to the influences of the school training, it was hypothesized that there would be similar changes in the scores for all variables of the control group, but that they would be less than those of the experimental group.

Because of the type of data yielded by the experiment, parametric methods did not seem suitable to evaluate the significance of these changes. Furthermore, on account of the "milieu" therapeutic effect, the scores of the control group were used as a base for evaluating the changes due to therapy in scores of the experimental group. Accordingly, three non-parametric methods were used to evaluate the results: (*a*) χ^2 test to compare the frequency of the gains and losses on the TAT for both groups; (*b*) a Mann-Whitney test to take into account the ordinal scores of these gains and losses from these two independent samples; and (*c*) a Wilcoxon matched-

pairs signed rank test to evaluate the changes within each group. The χ^2 test was corrected for continuity and the Mann-Whitney test was corrected for ties. As the direction of the changes was hypothesized for each trait, one-tailed tests of significance with a confidence limit of $P = .05$ were used.

The χ^2 test and the Mann-Whitney test served to evaluate the effect of group therapy over and above that of the "milieu" therapy. The Wilcoxon test served to determine the significance of changes within each group. Obviously, if the control group upon retest changed significantly, confidence in the significance of changes in the experimental group is lessened.

Since actual score changes on the TAT from first to second test are not large and the confidence level for significance is set only at $P = .05$, the following conditions were laid down: (*a*) based upon the Wilcoxon test, significant changes must occur from test to retest within the experimental, but not within the control group; (*b*) in the intergroup comparisons, significant differences must be found by both the χ^2 test and the Mann-Whitney test.

Results

Analysis of the data from the four sections of the experimental group yielded essentially similar results. Consequently, the results from the four sections of the experimental group were pooled. Table III lists the changes in mean scores from Test I to Test II for both experimental and

TABLE III

Changes in Mean Scores upon Repetition of TAT

Group	*N*	Hostility	Achievement	Blandness	Insecurity	Guilt
Experimental	43	−5.77	0.56	8.75	−2.04	−3.35
Control	43	1.26	−3.80	5.11	−6.04	2.00

control groups in all five categories of the modified TAT. A minus sign indicates that the mean score for Test II is less than that for Test I. It may be noted that the changes in mean score are small, and that they are always in the postulated direction for the experimental group, but that the control group has changes in the opposite direction for Hostility and Achievement.

In Table IV are listed the number of subjects whose scores in the first test were increased, diminished, or unchanged in the second test. Most of the experimental subjects diminished their Hostility scores upon retest; 6 of the 11 Ss who did not do so had been judged by the staff and by the therapist in the initial interview to have covert hostility. One might, therefore, expect that therapy would increase their overt hostility, which would be reflected in larger Hostility scores upon retest. These six subjects had an average increase in Hostility scores of 33.4 and a diminution in Achievement scores of 19.7; for all other subjects in the experimental

TABLE IV

DISTRIBUTION OF GAINS AND LOSSES UPON RETEST

	Hostility		Achievement		Blandness		Insecurity		Guilt	
Group	E	C	E	C	E	C	E	C	E	C
Gain	11	21	24	17	30	21	18	16	15	20
Loss	30	19	17	25	12	19	20	23	27	20
Same	2	3	1	1	1	3	5	4	1	3

E = Experimental group; C = Control group.

TABLE V

SIGNIFICANCE LEVEL OF NON-PARAMETRIC TESTS

Test	Hostility	Achievement	Blandness	Insecurity	Guilt
Chi-squared	.02	.08	.07	.38	.14
Mann-Whitney	.04	.16	.11	.15	.03
Wilcoxon					
Experimental	.02	.23	.003	.17	.11
Control	.39	.15	.09	.03	.22

group whose Hostility score diminished upon retest, the average change in score for these two categories was in the opposite direction and was 16.9 and 14.4 respectively. Such large reversals of scores for these six subjects confirms the impression that the overt expression of their hostility was initially inhibited.

The distributions of gains and losses from which the χ^2's were computed are given in Table IV and the corresponding P values are given in Table V. Table V also gives P values for the Mann-Whitney and Wilcoxon tests.

Based upon the conditions laid down above, it can be seen from Table V that the reduction in Hostility subsequent to therapy reached an acceptable level of significance, ($P=.05$) for all three non-parametric tests. The trend for the other traits for the experimental group is in the hypothesized direction (see Table IV) but does not reach an acceptable level of significance for all three tests. The changes for the control group for all traits except Insecurity do not reach an acceptable level of significance, nor was the direction of these changes adequately predicted. In other terms, our major hypothesis was confirmed, namely, that no such reduction in score takes place for those who have had no therapy. Changes in scores for other traits take place in the postulated direction but they are not such as to permit the rejection of the null hypothesis.

In general, there seems fair justification for the finding that even brief group psychotherapy produces significant changes in some objectively measurable responses of delinquent boys.

Summary

An experimental group of 43 delinquent boys was given brief group therapy with the object of reducing their Hostility. A matched control group, equal in size, received no therapy during the three and one-half week period of this experiment, other than the ordinary "milieu" therapy of the Training School. Both groups were given a modified TAT at the beginning and at the end of the experiment.

The major hypothesis that the experimental group would show a reduction of Hostility, as measured by the TAT, was confirmed within the 5 per cent level of significance.

Changes in the hypothesized direction of the TAT scores for the other four variables—Need for Achievement, Blandness, Insecurity and Guilt—occurred, but did not reach an acceptable level of significance.

REFERENCES

1. Burchard, Edward, M. L., Michaels, Joseph J., & Katkow, Benjamin. Criteria for the evaluation of group therapy. *Psychosomatic Medicine*, 1948, **45**, 257–74.
2. Frank, Jerome D. Some problems of research in group psychotherapy. *Int. J. Group Psychotherapy*, 1951, **1**, 78–81.
3. Gerston, Charles. An experimental evaluation of group therapy with juvenile delinquents. *Int. J. Group Psychotherapy*, 1951, **1**, 311–18.
4. Jones, Francis D., & Peters, Henry N. An experimental evaluation of group psychotherapy. *J. abnorm. soc. Psychol.*, 1952, **47**, 345–53.
5. Muench, George. Evaluation of non-directive psychotherapy by means of the Rorschach and other indices. *J. App. Psychology Mon.*, 1947, **13**.
6. Teuber, Hans-Lukas, & Powers, Edwin. Evaluating therapy in a delinquent program. *Psychiatric Treatment*, 1951, **31**, 138–49.
7. Zubin, Joseph. Design for the evaluation of therapy. *Psychiatric Treatment*, 1951, **31**, 10–15.

ASSOCIATIVE CLUSTERING IN THE RECALL OF MINIMALLY MEANINGFUL GEOMETRIC DESIGNS[1]

W. A. BOUSFIELD, H. BERKOWITZ, AND G. A. WHITMARSH

Department of Psychology, University of Connecticut

ASSOCIATIVE CLUSTERING may be defined as the tendency of subjects to recall the members of a randomized stimulus list in sequences of related items. Studies of this phenomenon have typically employed lists of words comprising subgroups or categories of two or more words each, with the words in each subgroup having been selected so as to be related to each other in a specifiable way. The types of relationship employed are evident in the following examples: taxonomic categories, Bousfield (**1**); *S-R* pairs selected from free associational norms, Jenkins and Russell (**6**); synonyms, Cofer (**3**); words in taxonomic categories with adjectival modifiers, Gonzales and Cofer (**5**). The effects obtained in these and related studies derive in large measure from previously established verbal habits of relatively long duration. The present study was designed with the purpose of reducing the effects of previous learning by using relatively novel stimulus items. To meet these specifications we chose non-verbal and minimally meaningful items as represented by geometric designs. In so far as definitive categories of such items could be prepared, it was reasonable to suppose that the organizational habits responsible for clustering would of necessity be developed primarily during the experiment, and that certain characteristics of this development, not evident from the use of categorized words, could be examined.

Though this study was designed as an investigation of associative clustering, it is apparent that it also deals with concept formation. The items chosen for each of the categories were similar in that they had common characteristics which were conceptual in nature. In these terms, it appeared that the experiment should provide two types of evidence of concept formation. The occurrence of any item in recall which met the specifications of a given category, but was not one of the stimulus items, could be regarded as evidence of the learning of the concept represented by the items in this category. The occurrence of clustering at a level beyond chance expectation would indicate not only the learning of the

[1]This paper is based on Technical Report no. 28 under Contract Nonr-631 (00) between the Office of Naval Research and the University of Connecticut.

concepts but also that the concepts were being used to organize the recall.

METHOD

Subjects

The Ss were 25 undergraduate students who were paid for their participation in the experiment.

Stimulus Materials and Apparatus

The 4 categories of stimulus items appear in Figure 1. As may be seen, the designs were made up as follows: Category A, a continuous line of variable direction, but

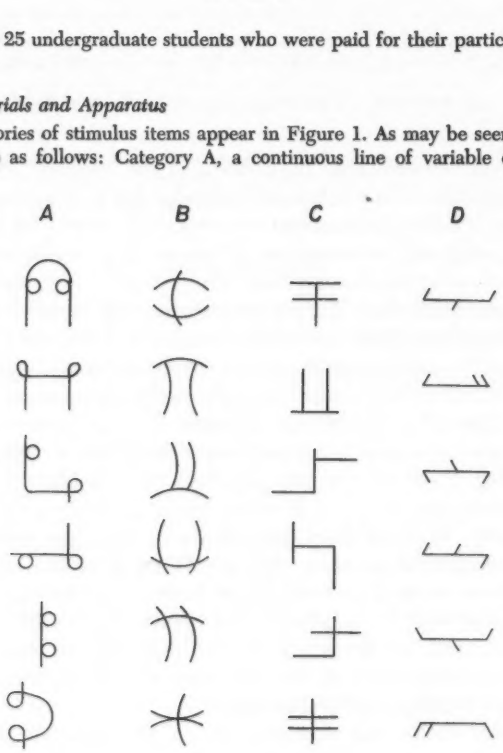

FIGURE 1

CATEGORIES OF DESIGNS

FIGURE 1. Categories of designs.

including 3 segments separated by 2 loops; Category B, 3 separate curved lines of equal length; Category C, 3 separate straight lines of equal length; Category D, a horizontal line to which are connected 3 separate tangential lines of equal length. These designs were selected from prepared pools of alternatives on the basis of the authors' agreement that they were minimally meaningful. Each design was then copied

on a 3 × 5 in. white card. To expose the designs to the Ss, the *E* used a 2 × 3 ft. screen containing a 1¼ × 2¼ in. opening behind which was a rack for holding the cards.

Procedure

Each S was run individually. In conducting the experiment, *E* informed S that a series of designs would be shown one at a time in the opening of the screen, and that the task was divided into 2 parts. First, S was to copy each design as it was presented. Second, after all the designs had been exposed and copied, S was to draw as many of the designs as he could. In doing this he was to draw the designs in a column on the blank sheet that had been given him, "*in the order in which they occur to you in your memory.*" The *E* also stated that a series of such trials for learning and recall would be given.

In conducting the experiment, *E* allowed about 7 secs. for S to draw each stimulus item as it was presented. The S started his recall immediately after the presentation of the final item, and continued for a period of 3 mins. The experiment allowed for a total of 5 successive learning-recall trials. In none of the recalls was S able to refer to his previous drawings. Prior to each learning trial, *E* randomized the stimulus items by shuffling the cards on which they appeared.

The 5 recall sequences of each of the 25 Ss were separately analysed according to the specifications of Cohen, Sakoda, and Bousfield (4) for the determination of the following scores:

1. *Number of stimulus items correctly recalled.* Since the stimulus lists comprised 4 categories of items with 6 items in each category the maximum attainable score was 24.

2. *Number of categorical intrusions.* A categorical intrusion was defined as an item having the actual components specified for a given category, but so arranged as not to match any given design. Such items were interpreted as indicating that S had learned the category rather than its specific members.

3. *Number of irrelevant intrusions.* When an item was not correctly recalled and could not be classed as a categorical intrusion, it was labelled an irrelevant intrusion.

4. *Number of repetitions, R.* Any sequence of two items which could be classed as belonging in the same category was labelled a repetition. In these terms, the number of repetitions in a cluster of 2 or more items was always one less than the number of items in the cluster.

5. *Ratio of repetition, RR.* This is the number of repetitions in a recall sequence divided by one less than the total number of items of all types appearing in the recall sequence. The measure has the advantage of simplicity and of taking extent of recall into consideration. Since its distribution tends to be approximately normal it is well adapted to tests of significance.

6. *Percentages of Ss showing Rs at various levels of significance.* Knowing the chance probability of a repetition in recall, the binomial distribution may be used as a basis of estimating the significance of *R* in relation to total recall. The proportion of Ss achieving significant clustering serves as a measure of the strength of the tendency to cluster in groups of Ss.

7. *Measures applied to the separate categories of items.* In order to compare the representatives of the 4 categories of items in recall, the following counts were made of the items classifiable as representing each category: (*a*) items correctly recalled; (*b*) number of categorical intrusions; (*c*) number of repetitions.

Results

Table I presents a summary of the data deriving from the first six of the measures which are general in nature. These findings warrant the following interpretations. Though the learning required was difficult, there was a fairly consistent progression in the correct recall of the

TABLE I

Summary of Data Based on Application of General Measures

Measures	Trials				
	1	2	3	4	5
Mean no. stimulus items recalled	5.76	8.56	10.36	12.36	13.72
Mean no. categorical intrusions	4.32	4.56	5.20	4.96	4.72
Mean no. irrelevant intrusions	.84	1.00	.68	.60	.32
Mean *R*	2.20	5.16	7.44	9.12	10.80
Mean *RR*	.22	.41	.49	.55	.60
Percentage of *Ss* clustering at .05 level	16.00	48.00	72.00	92.00	88.00
Percentage of *Ss* clustering at .01 level	0.00	20.00	60.00	68.00	68.00
Percentage of *Ss* clustering at .001 level	0.00	8.00	40.00	44.00	68.00

designs. The proportions of items classed as categorical intrusions appear to differ in two respects from the proportions of such items found by Bousfield and Cohen (2) in an experiment making use of four categories of meaningful words. In the present study these intrusions occurred at a relatively high frequency and they did not decline during the course of the learning. The presence of these intrusions as early as the first trial may be interpreted as indicating that the subjects were at this time making progress in learning the principles on which the categories were based. In comparing the number of designs classed as irrelevant intrusions with the number of such intrusions of words found by Bousfield and Cohen, we would submit only that these errors in recall were more prone to occur as a consequence of the use of our more difficult stimulus items. The number of repetitions, *R*, increased as a function of the number of trials, and also as may be noted, as a function of the number of items correctly recalled. The ratio of repetition, *RR*, is the preferred measure of the organization of recall relative to the categories of the stimulus items because it assesses clustering in relation to the amount of recall. In terms of this measure, the mean amount of clustering for the first trial, as indicated by the mean *RR* of .22, was at the level of chance expectation. The chance *RR* for our conditions, according to the formula for its assessment (4) is .2173. It may be noted, however, that 16 per cent of the subjects clustered at the .05 level after the first trial whereas according

to chance only 5 per cent would have been expected. The findings indicate that as a consequence of the first trial, the subjects made progress in learning the categories of the stimulus items since they produced a relatively large number of categorical intrusions, and the number who clustered significantly, though small, exceeded chance expectation. The data indicate a fairly steady increase in the *RR* and also an increase in the number of subjects who showed a significant amount of clustering as a function of the number of trials. Thus the learning was effective in bringing about a progressive increase in the extent to which the recall was organized on the basis of the categories of the stimulus items.

In the treatment of the data, consideration was given to the relative contributions of the four separate categories of designs to the over-all amount of clustering. To make this assessment, the twenty-five recall sequences for each of the five trials were separately analysed to determine the number of repetitions of items in each of the four categories. Using a split-plot design which took into account the fact that the subjects were the same throughout, an analysis of variance was made of the twenty distributions of repetitions. The mainplot analysis indicated that the variance attributable to the separate categories was significant at the .05 level of confidence. At the same time, the subplot analysis of the variance attributable to the five trials was significant at better than the .001 level of confidence. Thus it was evident that though there were significant differences in the contributions of the separate categories to clustering as measured by the incidence of repetitions, the increases in these scores attributable to the learning were highly significant. In order to obtain additional information regarding the differences attributable to the separate categories, counts were also made of their representations in terms of number of items correctly recalled and of categorical intrusions. It was clear from inspection that Category D, comprising the elements of a constant horizontal line and three variable connected lines, was most poorly represented in each trial by items correctly recalled. At the same time, this category was consistently responsible for the largest number of categorical intrusions. It is evident that though the trends indicated by the over-all measures applied to the data were relatively consistent, there was an appreciable variability in the extent to which the various categories of items contributed to the over-all scores.

Discussion

The results of the experiment indicate the feasibility of the use of categories of minimally meaningful geometric designs for exploring the organizational processes responsible for associative clustering. In terms

of the mean ratio of repetition, which is the measure best representing the over-all tendency to cluster in recall, the first presentation of the stimulus items resulted in no more than a chance amount of organization of recall even though several of the subjects did cluster significantly. The fact that nearly 40 per cent of the items in the first recall were categorical intrusions justifies the following interpretation of the nature of the learning process. It would appear that the subject learned the nature of the categories before they could serve as an effective basis for the organization of recall. Though the number of these intrusions appeared not to vary significantly with succeeding trials they constituted a progressively decreasing proportion of the number of items recalled. An analysis of variance of the number of repetitions in each of the four categories of items produced in the five successive recalls revealed significant differences attributable to the categories. In other words, the categories of items varied in difficulty in terms of the learning they required.

As was pointed out in our introduction, the present experiment may be regarded as dealing with concept formation, the concepts being the principles separately represented in the categories. The process of learning the concepts was that of identifying groups of perceptual elements when they varied spatially to form relatively novel patterns. The success of the subjects in learning the concepts was first indicated by categorical intrusions which were patterns containing the correct elements but which were combined to produce incorrect patterns. As the learning continued the concepts became increasingly potent for organizing recall. If these interpretations are correct, the procedures for inducing clustering should provide sensitive measures of the process involved in certain types of concept formation.

It may be submitted further that the method of inducing associative clustering by means of categories of minimally meaningful and relatively culture-free stimulus items should provide useful information on the effects of different types of mental impairment in organizational processes. The advantage lies in the fact that the organizational habits required for clustering in recall must be developed primarily during the experimental period, and the course of their development may be followed when successive trials for learning are given.

Summary

Associative clustering is defined as the tendency to recall the components of a randomized stimulus list in sequences of related items. This study reports an analysis of the development of associative clustering in an experiment making use of a list of 24 minimally meaningful and simple geometric designs divided into 4 distinctive

categories each of which was represented by 6 items. During each of 5 successive trials, Ss individually received a randomized presentation of the designs which they copied on paper. Immediately after each presentation the Ss undertook to reproduce as many of the items as they could in free recall. The analyses of the data indicated that though the mean amount of clustering for the first trial was at a chance level, several of the Ss clustered their recall to a significant extent. After the first trial, the mean amount of clustering and the number of Ss clustering significantly increased progressively as a function of the number of trials. Special attention is given to the occurrence of so-called categorical intrusions in recall. These are items meeting the specifications of the categories though not present in the stimulus list. Their relatively high incidence in the first recall suggests that the Ss tended to learn the categories before they used them as a basis for organizing their recall. The authors submit that the method of inducing associative clustering is applicable to the study of concept formation, and that the use of minimally meaningful and relatively culture-free items should provide a basis for studying certain mental abnormalities.

REFERENCES

1. Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. *J. gen. Psychol.*, 1953, **49**, 229–240.
2. Bousfield, W. A., & Cohen, B. H. The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. *J. Psychol.*, 1953, **36**, 67–81.
3. Cofer, C. N. A study in free recall based on synonyms. *Tech. Rep. no. 2*, ONR Contract Nonr–595 (04), 1956.
4. Cohen, B. H., Sakoda, J. M., & Bousfield, W. A. The statistical analysis of the incidence of clustering in the recall of randomly arranged associates. *Tech. Rep. no. 10*, ONR Contract Nonr–631 (00), 1954.
5. Gonzales, R. C., & Cofer, C. N. The effects of modifiers on the organization of verbal material as shown in clustering in free recall. *Tech. Rep. no. 6*, ONR Contract Nonr–595 (04), 1955.
6. Jenkins, J. J., & Russell, W. A. Associative clustering during recall. *J. abnorm. soc. Psychol.*, 1952, **47**, 818–821.

FURTHER COMMENT ON CLASSICAL AND INSTRUMENTAL CONDITIONING

MITCHELL M. BERKUN

HumRRO, Presidio of Monterey, Calif.

HEBB'S INTERESTING DISCUSSION (2) approaches but skirts an intriguing relation between instrumental and classical conditioning often considered but not yet explicitly stated. The Pavlovian procedure of using UCS to elicit a response which then can occur to CS is considered by Hebb to be actually an example of instrumental learning. That is, UCS becomes a reinforcement as a direct result of the response; therefore, the response is acquired by reinforcement. It appears only a small extension to consider UCS as a discriminative stimulus (S^D) in the Skinnerian sense (1), that is, a signal that 100 per cent of the responses will be reinforced. The CS then also acquires this discriminative property. The absence of CS is S^Δ, that is, a signal that the response will not be reinforced. Thus we have an example of control of a response by a discriminative stimulus, with response level high under S^D, low under S^Δ. The rules for operant or instrumental conditioning apparently then can be applied as well to classical conditioning, by combining Hebb's (2) and Schlosberg's (3) principle with that stated here. There may still, of course, be distinctions between reinforcing events that do or do not require drive reduction, such distinctions permitting two-factor theorists to stay in business. This distinction will apply when considering aversive conditioning. But a heuristic advantage may lie in considering that the principle of control of behaviour by discriminative stimuli and by varying the schedule of reinforcement may apply equally well to both classical and instrumental conditioning situations.

REFERENCES

1. FERSTER, C. B., & SKINNER, B. F. *Schedules of reinforcement.* New York: Appleton-Century-Crofts, 1957.
2. HEBB, D. O. The distinction between "classical" and "instrumental." *Canad. J. Psychol.*, 1956, **10**, 165–166.
3. SCHLOSBERG, H. The relationship between success and the laws of conditioning. *Psychol. Rev.* 1937, **44**, 379–394.

REVIEW ARTICLE

Motivation: A Systematic Reinterpretation. By DALBIR BINDRA. New York: Ronald Press Co., 1959. Pp. 361. $5.50.

FOR A LONG TIME psychology has had a series of convenient categories with which to subdivide its subject matter—learning, perception, motivation, and so on. These have served it well, but it has now become clear that the categories interact, that we cannot classify phenomena in this way and devise distinct explanatory hypotheses for each area. In particular, the study of "motivation" as a distinct aspect of psychology has led to the postulation of special instincts, drives, motives and so on to explain the phenomena it comprises. Lashley's warning that the concept of "drive" carried many of the disadvantages of the "instinct" which it displaced did little to stem the spate of "drive theories," of varying degrees of sophistication, which continues to the present day. Recently, however, discontent has been growing—Lashley's views have been re-emphasized, the impossibility of deriving all behaviour from a handful of primary drives has been recognized by all but a few stalwarts, and the oversimplification which unitary drive concepts entail has been revealed in a variety of ways. Furthermore, it is becoming clear that "drive reduction" cannot play quite the role in learning that has been assigned to it.

This, then, is a fitting moment to reconsider the problem of motivation. The basis for a new approach was defined by Hebb who, emphasizing that the brain is continually active, put the problem of motivation not as "Why does the organism behave?" but as "Why does it do this rather than that?" In this book Professor Bindra sets out with two problems—the processes involved in the development of the various activities in an organism's repertoire, and the factors controlling their appearance at any particular time and place. In discussing them, he is concerned primarily with those activities—hunger, thirst, sex, exploration, etc.—which are commonly studied in experiments on motivation, and he reviews the factors influencing their occurrence, strength, and integration. His approach is a fresh one, and he is thus able to present a reinterpretation of old experiments no longer hamstrung by the old explanatory concepts. He brings together material from all fields of psychology previously regarded as distinct, and thus sets the stage for a new synthesis.

After the introductory chapters the book falls into two main parts, the first of which is concerned with the development of motivational activities in the repertoire of the individual animal. Here Bindra's main thesis is

CANAD. J. PSYCHOL., 1959, 13(4)

that those features of behaviour which are referred to as "goal direction" are acquired through previous experience. While admitting the possibility of innate connections between certain stimuli and non-directed responses, such as reflexes, he denies the possibility of any such connection with directed responses: all aspects of goal direction are attributed to prior reinforcement by a given goal of responses that were contiguous with the goal. In support of this he reviews the literature on hoarding, eating and drinking, exploration, withdrawal and aggression, sexual and maternal behaviour, and so on. This provides a valuable corrective to those who would explain the organization of complex goal-directed activities as being purely instinctive or innate, and thus not requiring further analysis. Further, it throws into relief another difficulty common to many instinct theories, namely the classification of the organism's activities into quite separate compartments—a procedure which blinds the investigator to the influence of learning in one context on later behaviour in an apparently different one.

The problem of the development of behaviour is, however, one in which distinct questions are sometimes confused. All experiments in which the animal is subjected to an abnormal environment during development, and then tested for a given activity, give evidence relevant to the question "Are certain forms of previous experience essential to the development of the activity in question?" The answer to this is often "Yes." There remains, however, the further question "Can a *sufficient* explanation of the development of each goal-directed activity be found in previous individual learning?" Here, one feels, a more balanced approach could have been achieved if the evidence had not been drawn only from mammals. In the other vertebrate groups, at any rate, it is not possible to account for all cases of aggression or withdrawal solely in terms of responses either to conditioned stimuli or to strange or unexpected patterns: some strange patterns are more effective than others, and the more effective ones are different for each species. Similarly, the elicitation of sexual behaviour (rather than, say, fleeing, aggression, parental or social behaviour) by a dummy in a certain posture cannot be accounted for solely by previous social experience, even though such experience is necessary. In reacting, quite rightly, against the too easy explanations of instinct theorists, Professor Bindra lays little emphasis on evidence of this kind. His approach is unusually stimulating and fertile, but may prove one-sided.

The second half of the book is concerned with the factors determining the occurrence of the various activities. In successive chapters the author deals with habit-strength, sensory cues, arousal, and blood chemistry, providing in each case a stimulating selective review of the literature. In doing so he brings to bear on problems of "motivation" knowledge

from diverse disciplines—learning theory, perception, neurophysiology, endocrinology, and so on—so that the study of "Why the animal does this rather than that" is no longer confined by the narrow range of explanatory concepts previously used in this context.

Particular attention is paid to the question of arousal. Bindra emphasizes the loose nature of this concept, and employs it only in a rather general way. At the same time one may well feel that he uses it as an explanatory device in contexts in which it does not really help. Thus to say that differences in the effectiveness of a green cube and red cylinder in imprinting experiments is due to differences in their capacity to increase arousal is both untested and unhelpful. As another example he quotes an experiment by Hunt and Brady showing that a clicker, previously associated with an electric shock, causes a reduction in bar pressing by rats: he ascribes this to an increase in arousal beyond the optimum level, without consideration of the competing responses which the clicker presumably induces. Similarly, the role of androgen in increasing dominance is interpreted as a consequence of increased general arousal, with little regard to the fact that androgen affects certain types of behaviour more than others. Furthermore, the effect of androgen on sexual behaviour is ascribed largely to its effect on dominance: such a view could have only a limited range of applicability, for in many species the male becomes subordinate to the female at mating. However, it is right that the explanatory possibilities of a fashionable concept should be explored, and the fact that it is overstretched in a few contexts scarcely lessens the value of this review and resynthesis of the data bearing on it.

In the final chapter Bindra draws together some of the earlier conclusions and provides an interesting discussion of "functional autonomy." The apparently self-reinforcing nature of many activities has long been a puzzle, and the labelling of them as autonomous has hardly helped us in understanding them. Bindra, however, points out that such activities are autonomous only with respect to a limited set of conditions, and discusses them in the light of principles known to be important in other contexts—substitution of cues, increased effectiveness of cues, reinforcer substitution—and thus makes autonomy respectable. This will have wide implications, especially for psychopathologists.

This book, then, represents an important step forward. The two issues taken up in this review (the emphasis on reinforcement in the first part of the book and arousal in the second) do not detract from the main theme: the fresh approach will do much to unshackle psychology from concepts which have outlived their usefulness, and indicates many of the lines which research will take in the next decade.

R. A. HINDE

Cambridge University

Personality Patterns of Psychiatrists: A Study of Methods for Selecting Residents. Vol. 1. By ROBERT R. HOLT and LESTER LUBORSKY. New York: Basic Books, Inc., 1958. (Menninger Clinic Monograph Series no. 13). Pp. xii, 386. $7.50.

IN A SENSE the title of this book is a misleading one. The subtitle "A Method for Selecting Residents" is more accurate, for it is the story of the selection of psychiatric residents at the Menninger School of Psychiatry, as carried out over a 10-year period following the war.

This book and its companion volume, which provides information on details of research methods and quantitative tables, is the story not of an organized experimental design established to verify the validity of the methods used; rather it describes an *a priori* scheme with certain residents being selected and others rejected on the basis of the previously established concepts. Thus there is provided no basis for any real definitive evaluation of the procedure used. In essence the study was the by-product of a procedure considered essential for the running of the school, subject to all the pressures of such a procedure. The authors have undoubtedly made the best of such an arrangement, established for purposes other than research on methods of selection. This is the story primarily of those selected, with efforts being made to relate subsequent experience of the residents with initial findings.

The selection was carried out partly on the basis of a psychiatric interview—heavily biased toward analytical concepts—and partly by various psychological tests. The weighting of the two methods of selection is not made clear. Extensive use was made of various statistical techniques in validating the procedures.

The final concept of the ideal psychiatrist is a rather disappointing one. One reads the book looking forward to this aspect with the hope that one will obtain insight into the personality structure of a group that has been described as differing from other physicians. Instead of this, one obtains a picture of what has been called the "ideal man"—certainly a very select type of individual.

The book is to be recommended for anyone interested in selection in general, as a description of one method of approach well told and documented. It is not recommended as a model in experimental design.

G. E. HOBBS, M.D.

The University of Western Ontario

An Introduction to the Behaviour of Invertebrates. By J. D. Carthy. London: Allen & Unwin Ltd., 1958. Pp. xvi, 380. 45*s*.

This is a useful book, and in some ways a good one. But its conception of the scope of animal behaviour is surprisingly narrow, and its title is therefore misleading. Even in its own chosen field its presentation of modern knowledge often lacks adequate perspective.

Of the twelve chapters that follow a three-page introduction, all but one are exclusively concerned with reception—the sense organs, the light senses, perception of polarized light, visual form and movement, reactions to heat, the position senses, the perception of vibration, the tactile senses, the chemical senses in general, the contact chemical senses, and humidity sensitivity. The result is a widely informed summary of recent research on the functions of invertebrate receptors, based on a bibliography of 600 items. As such it includes a considerable amount of conveniently classified information, for which the reader will be duly grateful.

The book might do as the review of a random collection of miscellaneous observations on invertebrate reception, but it is less satisfying as a general introduction to the subject as a whole; for the author's treatment shows no great sense of perspective, no awareness of the wider significance of the results obtained in the total field of animal form and function. In the case of insects, for instance, there is constant confusion between imaginal ocelli and larval stemmata (also here called ocelli). Nor are the very different functions of the various pairs of stemmata of, for instance, larval *Cincidela* considered, with their separate control of general orientation, the level of emergence from the burrow, and the closing of the jaws. Again, there is no appreciation of the very different possibilities of the single eye of *Cladocera*, the normal twin eyes of the majority of species, the triple eye of Copepods, the four, six, or eight compound eyes of spiders, or the scores and sometimes hundreds that are found among Lamellibranchs. Similar remarks could be made in the case of other senses. In general, the function of the individual sense organ has not been adequately reintegrated into the total setting of the animal's morphology and physiology. This is the result of the author's too exclusive interest in the functions of the senses. Reception is certainly important. But can all other aspects of animal behaviour be crammed into a single final chapter, especially if by "whole patterns of behaviour" one means no more than ill-defined reactions to combinations of simple stimuli? The part of this chapter devoted to food selection is finally reduced to one page, and the section on mate selection is shorter still. Both are introduced as examples of perception. In fact, sense reception has been turned into the whole of animal behaviour, and other subjects are neglected.

The central nervous system, for instance, is not discussed, though it would seem essential to understand the effect of its different concentration in various groups of molluscs, and of the relative development of the cerebropleural ganglia at one end and the visceroparietal at the other; or again to appreciate how much of the elaborate egg-laying behaviour of the silk-moth is organized by the terminal concentration of abdominal ganglia, and how little it depends on the insect's head. Centrally co-ordinated motor mechanisms are nowhere considered, and such distinctive operations as grooming, eating methods, or copulation techniques are therefore taken for granted. Kineses and taxes are mentioned in various chapters, but their modes of operation remain unexplained. No interest whatever is shown in the effector motor performance required to carry out directed movements, such as flying, swimming, crawling, and running; yet these also are species-specific modes of behaviour, derived from the co-ordinated activities of groups of cilia or muscles, controlled (in the latter) by the nervous system, based on a typical morphology, and stiffened by a peculiar exo-, hydro-, or haemo-skeleton. The same indifference is shown for the problem of animal constructions: crabs' burrows and caterpillars' webs are mentioned incidentally; but there is no information as to their shape or orientation or method of construction. Nor is there any discussion of spiders' webs and cocoons, of the byssus of Lamellibranchs, of the nests and cells and galleries of insects or the cases and cocoons and chimneys of their larvae, of the tunnels of shrimps, the borings of bivalves, the excavations and mucous nets of worms, etc. Are we to believe that no behaviour is involved in all these remarkable building activities?

Social behaviour is likewise neglected. Courtship, for instance, is frequently mentioned, but always as a response to a specific stimulus. Neither the organized pattern of the individual's motor display nor the dynamic social encounter of the pair is ever studied for itself. (One wonders, in this connection, why both the visual display of Cephalopods and the spectacular tactile mating patterns of land pulmonates are deemed unworthy of notice.) Threat display is mentioned only once (for a spider), and is not further analysed; but dominance (as in crayfish) is completely ignored. The care of eggs or offspring is omitted; social co-ordination is given a passing glance only in relation to the dances of bees; migration is considered only once as an effect of light intensity on the vertical movements of plankton. Learning is occasionally referred to, but its theory and implications are dismissed with a reference to Thorpe's latest book—but surely it is essential to know that a given species (or family or order) is, in principle, capable of learning, what type of learning it is able to master, and how much learning it does in fact normally achieve.

The general ecological factors that make up an animal's specific environment are likewise ignored—salt water or fresh, running or stagnant, benthic or pelagic, neritic or abyssal, rocky or sandy or muddy bottoms. Nor is there any attempt to evaluate the relative importance of the various stimuli that impinge upon the animal's sense organs—surely a fundamental problem for the understanding of invertebrate behaviour.

An immense amount of information is available in the zoological literature on all these points; and obviously it would all be relevant to the scientific study of animal behaviour. No one would expect an Introduction to review all this material in a critical manner. But surely a general survey of the field as a whole should have been provided, with a balanced discussion of the principal factors involved and of the main problems to be solved. Moreover, since an Introduction cannot possibly claim to include all the available material, it should introduce the reader to the relevant secondary sources of essential information—Pelseneer's classic *Essai d'éthologie zoologique* for the molluscs, Bonnet's *Bibliographia araneorum*, both Wigglesworth and Roeder on insect physiology, Von Lengenkern on the care of offspring among Coleoptera, and other general works of similar scope. As an introduction to the behaviour of invertebrates, this book is therefore disappointing. As a somewhat uncritical digest of recent work on their sense organs, it is a useful addition on the specialist's bookshelf.

D. H. Salman

Université de Montréal

De l'actinie à l'homme. II, *De l'instinct animal au psychisme humain.* Par Henri Piéron. Paris: Presses Universitaires de France, 1959. Pp. 264. 1400 francs.

This is the second and final volume of Prof. Piéron's collected "studies on comparative psychophysiology." The book is divided into four parts. The first concerns the chemical senses: a general review of their evolution (1942), four papers on molluscs (1908–41), and four more on olfactory recognition among ants (1904–6). The second part contains a study on the notion of instinct (1914), two on the geotropism of snails (1910–29), four shorter ones on death-feigning in insects (1907–26), and several more on the autotomy of crabs (1907) and various forms of self-mutilation in insects (1907–26). Few if any of these problems have been re-examined since then with the same species, so that this early experimental work retains its fundamental usefulness. The third section includes two papers on the physiology of emotion (1907, 1923), one on sensory affectivity

(Mooseheart Symposium, 1951), and another on the neurophysiology of pain (Stockholm Congress, 1952). Of these, the last is no doubt the most important, as a comprehensive review of the subject in the light of the author's own research; and it would deserve more attention than it has received on this side of the Atlantic. The fourth and last part contains more general essays on the social factors in human development (1952), and the importance of Pavlovian conditioning in personality growth (1951). This collection of widely scattered papers will prove invaluable to specialists, and remain as a fitting memorial to the varied interests of one of the major psychologists of the early twentieth century.

D. H. SALMAN

Université de Montréal

The Security of Infants. By BETTY MARGARET FLINT. Toronto: University of Toronto Press, 1959. Pp. x, 134. $3.50.

THIS BOOK is an important addition to the series of studies which have been inspired by W. E. Blatz's concept of security and published by persons associated with the Institute of Child Study in Toronto. Mrs. Flint's primary purpose is to report a pioneering attempt to develop a scale for assessing objectively the mental health of children up to 24 months of age. She also provides a welcome clarification of Blatz's theory of personality development at the infant level.

The Infant Security Scale consists of a diagnostic check-list of items describing behaviour characteristic of babies. Four scales are provided, one for each half of the first two years of life. The items were selected, in terms of the security concept, from lists of behaviour criteria derived from extensive observations of infants in clinics and in their own and foster homes. The items are classified under the headings: secure, insecure, and deputy agent. Infant security, or mental health, is defined as a serene state of mind arising from a feeling of self-worth and a conviction that the world is benign, and is thought to be reflected in a willingness to accept care and an eagerness to put forth effort towards mastery of the environment. Hence, the scale attempts to measure the extent to which the child accepts dependence and takes advantage of opportunities for independent effort in a variety of situations such as feeding, toileting, play, and social activities.

Infant security is assumed to arise through a close dependent relationship with a mothering figure. If the mothering agent is dependable, that is, anticipates and meets the child's needs in a comfortable and consistent way, the child is said to develop dependent trust or a basic sense of

trust in his world; if she is "warm" in her feelings toward him and expresses this warmth through her ministrations during care, the child is said gradually to feel himself to be a person of value and to develop self-trust. This latter feeling is further enhanced by rewarding self-initiated effortful activity, especially when this receives maternal approval. Poor mental health is thought to result from apprehension and uncertainty instilled by the unpredictable mother whose care of the child is based on her own whims and goals rather than the baby's needs. The insecure child, or secure child under stress, may regress or employ a number of defensive techniques referred to as deputy agents.

The experimentalist will find it easy to criticize the research reported here with respect not only to the statistical treatment of the data, but also to certain aspects of the procedures employed. Two studies were undertaken in an attempt to validate the scale. In one a comparison was made of the scores of twenty "well-adjusted" infants, living at home and cared for by their own mothers, with those of thirteen "poorly-adjusted" infants in foster homes, who were reported by the supervising agency to have made unsatisfactory adjustments to these homes. The "well-adjusted" obtained higher mean and median security scores than did the "poorly adjusted" children, but the groups differed widely in both socio-economic status and developmental level, the differences in both cases favouring the "well-adjusted" group. More serious still is the fact that, in the case of the "well-adjusted" children, endorsement of items was done, independently, by the babies' own mothers, whereas scores were assigned for the "poorly adjusted" infants by the experimenter in co-operation with the foster mothers. It is hard to discount the possibility of maternal bias, especially in view of the fact that these mothers were, apparently, not only aware that the scale purported to measure mental health but were also, for the most part, sophisticated with respect to the security concept. The second study reports changes in scale scores of sixteen institutionalized babies over a period of approximately five months. While there were large individual differences in the scores of these infants, all but one showed some decline in score over the period during which they were observed. Here again, the absence of a suitable control group makes it difficult to evaluate these findings.

The limitations of this work are offset by the fact that it is published as an interim report. The author recognizes that the validity and reliability of the scale have not, as yet, been satisfactorily established. In fact, it should be noted that a revision of the scale presented here has already been made and is available in mimeographed form. Modifications of the present scoring systems, which do not take into account performance on items classified under the heading insecure, are anticipated.

Mrs. Flint is to be congratulated for making this first attempt to provide a clinical instrument of a sort which is not now available, and even in its present form the scale should prove of practical value to clinicians and social workers responsible for the welfare of young children. The fact that the scale has been constructed in terms of a specific body of theory should increase its probable usefulness as a tool for research.

MARY J. WRIGHT

University of Western Ontario

Group Processes. By BERTRAM SCHAFFNER. New York: Josiah Macy Foundation, 1959. Pp. viii, 266. $4.50.

CE LIVRE, édité par Bertram Schaffner, contient à la fois les exposés et le compte rendu des discussions qui ont suivi de la "Fourth Conference on Group Processes," tenue à Princeton, N.J., les 13, 14, 15 et 16 octobre 1957. Ces rencontres sont organisées et financées par la Josiah Macy, Jr. Foundation.

Les trois précédentes conférences sur les processus de groupes avaient surtout porté sur les phénomènes de groupes, observés chez les animaux et les jeunes enfants. Cette quatrième conférence essaie de dresser le bilan des données acquises sur les processus de groupes d'enfants plus âgés. Cette conférence s'ouvrit par un hommage à la mémoire du Dr. Frieda Fromm-Reichmann auquel participèrent: Gregory Bateson, Ray L. Birdwhistell, et Frank Fremont-Smith. Suivirent quatre exposés: deux de Fritz Redl, un d'Alex Bavelas et le quatrième de Konrad Z. Lorenz.

Le premier exposé de Fritz Redl portait sur les répercussions psychologiques de la structuration des jeux d'enfants. Les observations sont rigoureusement obtenues et interprétées avec un excellent sens critique. Le lecteur s'étonnera peut-être de ne trouver, ni dans l'exposé, ni dans la discussion, aucune référence faite aux travaux si concluants de Jean Piaget sur précisément ces problèmes et ce qu'ils révèlent sur la genèse du sens social chez l'enfant. Dans son second exposé, Fritz Redl met en cause certaines théories de la personnalité en les confrontant avec ses propres observations sur la psychologie des groupes. Ce second exposé nous a de beaucoup plus satisfait que le premier. Redl s'y révèle le meilleur théoricien présentement de la psychologie des phénomènes de groupes chez les adolescents.

L'exposé d'Alex Bavelas porte sur des problèmes de communication, en montrant comment les interactions dans un groupe sont affectées par la dimension du groupe et les structures ambiantes. Quant à Konrad Z.

Lorenz il a su mettre en évidence, avec une grande rigueur scientifique, le rôle de l'agression dans la formation des groupes.

Le mérite de l'éditeur est très grand de nous présenter non seulement les exposés, mais les confrontations qu'ils ont suscitées chez les participants de ces rencontres. Ces échanges et ces discussions, loin d'ébranler, de miner ou de saboter les mérites et la valeur des exposés, offrent une chance exceptionnelle au conférencier de préciser sa pensée, de la nuancer, de l'étayer et au besoin de la clarifier. A condition que ces échanges se poursuivent comme le souhaitait Frieda Fromm-Reichmann "with tenderness and respect for the other person."

BERNARD MAILHIOT, O.P.

Université de Montréal

CANADIAN JOURNAL OF PSYCHOLOGY

VOL. 13, NO. 4, DEC., 1959

PUBLISHED FOR THE

CANADIAN PSYCHOLOGICAL ASSOCIATION BY THE

UNIVERSITY OF TORONTO PRESS

JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR

JEAB publishes original reports of experimental investigations relevant to behavioral processes in individual organisms. The interests of its editorial board include a wide range of areas in which the experimental analysis of behavior is appropriate, both in its own right, and in its application to Physiology, Pharmacology, Education, Psychiatry, Genetics, and other allied disciplines.

Issues appear at least quarterly. First issue, January 1958.

Subscription: A.P.A. Member, $9.50. Student, $6.50. Regular, $12.00.

Mail subscriptions to Dr. O. R. Lindsley, Behavior Research Laboratory, Metropolitan State Hospital, Waltham, Massachusetts, U.S.A.

Published by THE SOCIETY FOR THE EXPERIMENTAL ANALYSIS OF BEHAVIOR